Temporal Dynamics of Binocular Disparity Processing in the Central Visual Pathway

Michael D. Menz and Ralph D. Freeman

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

Menz, Michael D. and Ralph D. Freeman. Temporal dynamics of binocular disparity processing in the central visual pathway. *J Neurophysiol* 91: 1782–1793, 2004. First published December 10, 2003; 10.1152/jn.00571.2003. To solve the stereo correspondence problem (i.e., find the matching features of a visual scene in both eyes), it is advantageous to combine information across spatial scales. The details of how this is accomplished are not clear. Psychophysical studies and mathematical models have suggested various types of interactions across spatial scale, including coarse to fine, fine to coarse, averaging, and population coding. In this study, we investigate dynamic changes in disparity tuning of simple and complex cells in the cat's striate cortex over a short time span. We find that disparity frequency increases and disparity ranges decrease while optimal disparity remains constant, and this conforms to a coarse-to-fine mechanism. We explore the origin of this mechanism by examining the frequency and size dynamics exhibited by binocular simple cells and neurons in the lateral geniculate nucleus (LGN). The results suggest a strong role for a feed-forward mechanism, which could originate in the retina. However, we find that the dynamic changes seen in the disparity range of simple cells cannot be predicted from their left and right eye monocular receptive field (RF) size changes. This discrepancy suggests the possibility of a dynamic nonlinearity or disparity specific feedback that alters tuning or a combination of both mechanisms.

INTRODUCTION

Stereoscopic depth perception has been studied from theoretical, behavioral, and neurophysiological perspectives. Most behavior work has been aimed at establishing empirical parameters of stereoscopic function and performance levels. Early neurophysiological studies neglected mechanisms and assumed that cells with responses that changed with retinal disparity of the stimulus were depth detectors. More recent physiological work has included theoretical proposals that were tested experimentally. The result is a modified energy model that accounts for basic features of the neural mechanism of stereopsis (e.g., Freeman and Ohzawa 1990; Ohzawa et al. 1990; Qian and Andersen 1994; Qian et al. 1994).

An early theoretical proposal was concerned with the correspondence problem in stereopsis (Marr and Poggio 1979). This refers to the ambiguity of correspondence between left and right images that occurs from binocular viewing. The brain must choose the correct depth plane from several possible ones to process appropriate stereoscopic information. To solve this problem, coarse scale disparity matches were proposed to occur first. This would be followed by fine scale matches (Marr and Poggio 1979). In other words, a coarse-to-fine scaling process could be used to provide correct stereoscopic matching. Other theoretical

ideas envisioned similar coarse scale disparity matches that were followed by fine scale adjustments (Anderson and Van Essen 1987; Nishihara and Kimura 1987; Quam 1987).

These theoretical notions have been explored in behavioral studies. Sensitivity has been found to improve for line length, orientation, curvature and stereoscopic depth over a period of ≥ 1 s (Watt 1987). The interpretation of these findings may be made in terms of a coarse-to-fine temporal analysis of spatial features. Another study was aimed at determining if different spatial scales interact in stereopsis (Rohaly and Wilson 1993; Wilson et al. 1991). Diplopia threshold was determined at two separate spatial scales. Results suggested that coarse scale disparity processing constrains that of fine levels within a given range. Two experiments were performed in another study in which spatially filtered targets were used. In one, temporal sequences were shown in which full-bandwidth targets were compared with those in which selected frequencies were presented first. In a second experiment, a human face was used in a similar way. Results of both experiments provide clear evidence for anisotropic temporal processing. Specifically, the most efficient perceptual processing occurs when spatial information is presented temporally in a low-to-high spatial frequency sequence. (Parker et al. 1997)

Other psychophysical investigations of stereopsis suggest that there is also a fine-to-coarse process. In one study, an ambiguous coarse scale stimulus was presented that could be perceived with either crossed or uncrossed disparity. When a fine scale stimulus was added, the ambiguity at coarse scale was removed, suggesting a fine-to-coarse disambiguation process (Smallman 1995). However, this study also showed that a coarse scale stimulus could disambiguate that of a fine scale. Therefore results of the study support both coarse-to-fine and fine-to-coarse processes.

Considered together, results of the behavioral studies suggest that both coarse-to-fine and fine-to-coarse processes may apply to stereopsis. Surprisingly, until our recent study (Menz and Freeman 2003), no physiological data on this issue have been available. Neuronal temporal analysis on a fine time scale has become possible with relatively recent techniques such as reverse correlation analysis (DeAngelis et al. 1993a,b; Freeman and Ohzawa 1990; Jones and Palmer 1987). Results of orientation (Ringach et al. 1997) and spatial frequency (Bredfeldt and Ringach 2002) tuning studies using this technique show prominent changes as responses evolve over time. This type of information is important because it provides clues about neural circuitry. In ideal cases, for example, it may be

possible to obtain evidence consistent with feed-forward or feedback models of visual processing.

We have carried out a temporal analysis over a brief time scale (i.e., 40 ms) for a population of neurons in lateral geniculate nucleus (LGN) and visual cortex. In addition to determining specific temporal features relevant to stereoscopic processing, our aim was to obtain information concerning the theoretical proposal of a coarse-to-fine sequence as outlined in the preceding text. Our results provide clear evidence that is consistent with this hypothesis.

METHODS

Experiments are conducted using anesthetized, paralyzed cats. The equipment and procedures for surgery, animal maintenance, single-unit recording, receptive field (RF) mapping, and some data-analysis techniques have been described in detail in previous papers (e.g., Anzai et al. 1999a,b). The description provided here emphasizes specific procedures that are relevant to this study.

Surgical procedures and animal maintenance

Following standard preanesthetic procedures, isoflurane is used to anesthetize the animal. Femoral veins are cannulated, a tracheal tube is placed, and a craniotomy and durotomy are performed (H-C P4 L2 for visual cortex and A6 L8 for LGN). After surgery, thiopental is used to maintain anesthesia. Each animal is assessed individually to determine an adequate level of anesthesia, which generally ranges from ~1.5–2.5 mg $\cdot$ kg^{-1} $\cdot$ h^{-1}. After anesthesia level is stabilized over 1 h, a muscle relaxant (gallamine triethiodide, 10 mg $\cdot$ kg^{-1} $\cdot$ h^{-1}) is used to prevent eye movements during visual stimulation. Pupils are dilated, nictitating membranes are retracted, and contact lenses with 4-mm artificial pupils are positioned. A reversible direct ophthalmoscope is used to image the optic discs on a tangent screen to infer areae centrales locations (Bishop et al. 1962). Core body temperature, electroencephalogram (EEG), electrocardiogram (ECG), heart rate, intratracheal pressure, and expired CO_2 are all monitored continuously throughout each experiment.

Recording procedure

Visual stimuli are generated by a computer with two high-resolution graphics boards that runs custom software as described previously (DeAngelis et al. 1993a,b). To map cortical RFs, a dichoptic one-dimensional binary m-sequence noise stimulus is used (Anzai et al. 1999a,b). A monocular stimulus is used for LGN cells. Sixteen long adjacent bars are presented to each eye at optimal orientation (Fig. 1A). The width of the bars is approximately one-fourth the

period of the optimal frequency. This square pattern is centered over the RF. Each of the 16 bars is either bright or dark, and the background is the mean luminance of the bars.

Nonlinear binocular disparity tuning and monocular RFs are determined simultaneously. Each spike train is cross-correlated with the stimulus sequence by means of a fast m-transform (Sutter 1991) to obtain space-time RF maps (Fig. 1B). This is repeated for all time delays of interest (0–200 ms) in increments of 5 ms to obtain a space-time RF. A nonlinear binocular interaction map is obtained by
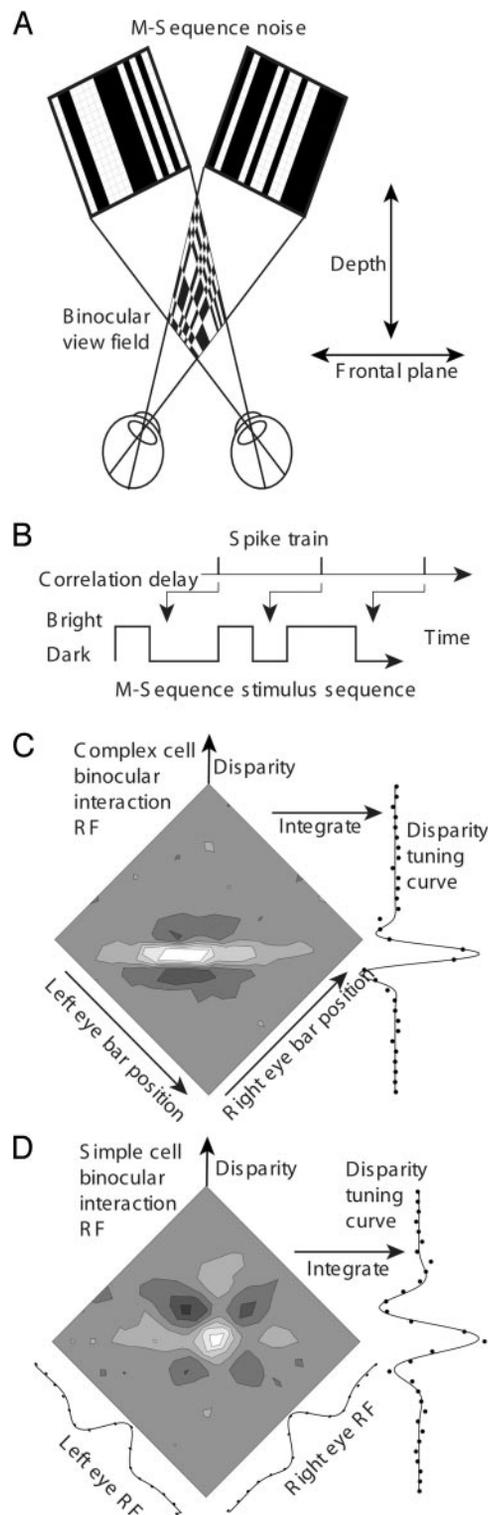

FIG. 1. Method for obtaining disparity-tuning curves and monocular receptive fields. A: dichoptic m-sequence dense noise stimulus presented at optimal orientation for each eye. The binocular view field is a representation of the binocular combination of stimuli with either bright bars or dark bars in both eyes (same contrast - white) or a bright bar in 1 eye and a dark bar in the other eye (opposite contrast - black), presented at various depths across a frontoparallel plane. B: the spike train is cross-correlated to the binocular combination of stimuli at various delays to obtain a nonlinear binocular interaction map at each correlation delay. The spike train is cross-correlated to each monocular stimulus to obtain the linear monocular receptive fields. C: an example of a binocular interaction map for a complex cell at the optimal time delay (i.e., strongest signal). This map describes the response of the cell to the same contrast (white) or opposite contrast (black) bars presented at various depths along a frontoparallel plane. This 2-dimensional (2-D) map is integrated along lines of constant disparity to obtain a disparity-tuning curve that is fit with a Gabor. D: simple cell example including the 1-D left and right eye receptive fields, which are also fit with Gabors. For more details concerning binocular m-sequence, see our previous papers (Anzai et al. 1999a,b).

cross-correlating the spike train with each spatial location in the binocular view field pattern. In the maps shown in Fig. 1, *C* and *D,* the white and dark regions indicate excitatory responses to same or opposite polarity bars, respectively.

Data analysis

Our analysis smoothes or interpolates the RFs within the 40-ms duration stimuli and uses eight bins to achieve space-time RFs with 5-ms interval correlation delays. It is possible that the autocorrelation of the stimulus produces an artifactual dynamic (Theunissen et al. 2001). To test for this possibility, we also generated RFs using bins equal in size to the stimulus duration (40 ms). The two-dimensional binocular interaction maps are reduced to one-dimensional disparity-tuning data by integrating along lines of equal disparity (Ohzawa et al. 1997) (see Fig. 1, *C* and *D*). These data are fit with a Gabor function by the Levenberg-Marquardt algorithm (Press et al. 1992)

$$S(d) = \exp(-(d - d_0)^2/2\sigma_s^2) \cos(2\pi f_{ds}(d - d_0) + \phi) \qquad (1)$$

where d is disparity, d_0 is the center position, σ_s is the width or size parameter, f_{ds} is the disparity frequency, and ϕ is the phase. For LGN cells, the fit in the spatial domain is a single Gaussian function

$$\text{LGN}(x) = \exp(-(x - x_0)^2/2\sigma_c^2) \qquad (2)$$

Where x is position, x_0 is the center position, and σ_c is the size parameter. A difference of Gaussians (DOG) fit was attempted, but the surround was so weak at nonoptimal time slices that this procedure was not practical. For the purpose of this study, a single Gaussian function is suitable.

The most direct, assumption-free method of analyzing frequency content is to take the Fourier Transform at each time delay and examine the change in optimal frequency and bandwidth. The frequency data were fit with a Gaussian function and the dynamics measured this way closely match the method of fitting a Gabor function in the spatial domain (data not shown).

At each time slice, the value of a parameter is normalized (i.e., divided) by the value at the optimal time slice. A linear regression is performed on the parameters as a function of time delay relative to optimal. The slope is used as a measure of the rate of change of the parameter. The value from the regression is multiplied by 1,000 yielding units of %/10 ms. For the monocular RFs of disparity-tuned simple cells, a separate regression is performed for the left and right eye findings, and the two numbers are averaged to obtain an overall monocular result. Unless otherwise noted, statistics are based on a standard normal distribution as described by the Central Limit theorem, which requires a large sample size.

RESULTS

We have analyzed responses of a total of 186 neurons in LGN (39) and striate cortex (147). Of the cortical cells, we classed 60 and 87 as simple and complex, respectively, using standard criteria (Hubel and Wiesel 1962; Skottun et al. 1991). Although LGN cells are monocular, except for modest dichoptic effects, we include them in this study to gain insights into possible mechanisms of organization of the neural system for disparity processing. Our main focus is an examination of temporal response characteristics of each cell type in the visual cortex. We examine both monocular and binocular properties of simple cells to compare temporal properties in the processing sequence. For LGN cells, we examine changes in RF size during responses. We use best-fitting Gaussian functions to make this assessment. For cortical cells, best-fitting Gabor functions are used and the variable of interest is correlation delay.

LGN

We first consider the spatial properties and temporal dynamics of LGN cells. We assume a serial processing system in which input is fed from LGN neurons to simple and then to complex cells. The actual nature of this system is not crucial for our analysis, and we assume that there is also parallel processing.

An example of temporal characteristics of the RF profile for an LGN cell is shown in Fig. 2. The data in *A–C* are fit with single Gaussian functions. They show, respectively, RF profiles for 15 ms before the optimal delay, at the optimal value, and 15 ms after
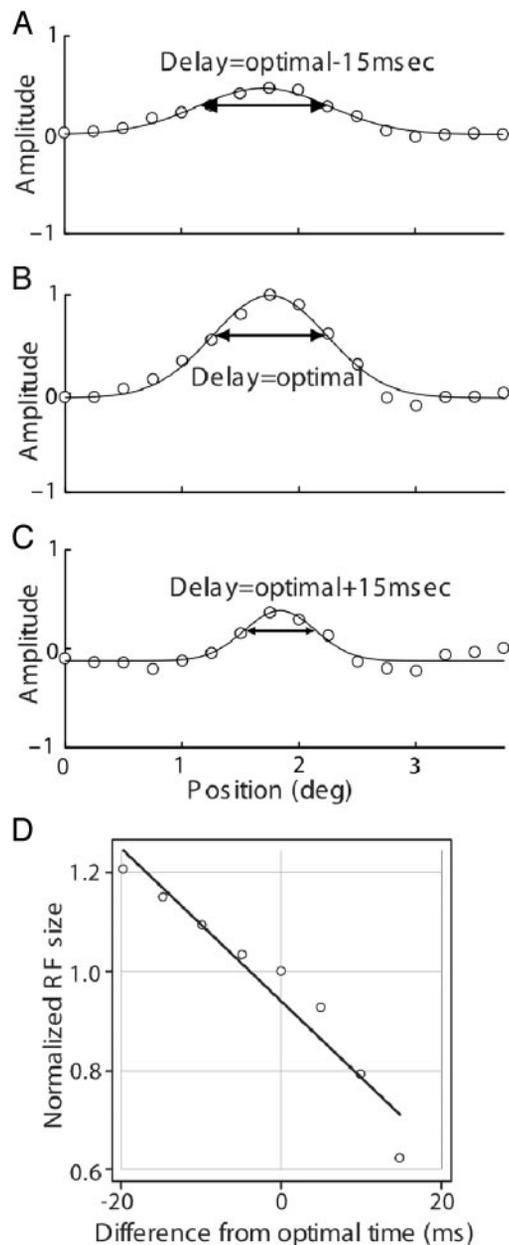
FIG. 2. Results for a lateral geniculate nucleus (LGN) cell demonstrate that the center size becomes smaller with increasing correlation delay. *A–C*: 1-dimensional receptive field fit with a single Gaussian (not a difference of Gaussians) at 15 ms before the optimal delay (*A*), optimal delay (*B*), and 15 ms after the optimal delay (*C*). ↔, widths at half-maximum amplitudes. *D*: the normalized size parameter of the Gaussian fit (●) is plotted at each correlation delay. Size decreases smoothly and extends beyond the optimal delay [thin regression line slope = −15.4 (%/10ms)].

optimal delay. Widths at half-maximum amplitudes are designated (↔). Normalized RF width as a function of time delay, given in Fig. 2*D,* shows a clear reduction in RF size with increasing correlation delay. For this cell, the center size decreases at a rate of 15.4%/10 ms. Note that the range of usable time delays is relatively small because the duration of the first phase of the temporal response is short. This is due to LGN preference for high temporal frequencies. Note also that in the example shown in Fig. 2, LGN center-surround organization is not evident in the early part of the response (*A*) but it appears weakly at a subsequent time slice (*C*). This result is consistent with our previous work showing that LGN surrounds are frequently time-delayed relative to the center response (Cai et al. 1997).

The temporal relation of the center and surround RF regions of LGN cell responses is the basis of dynamics at this stage of visual processing. If center and surround respond with the same temporal characteristics, the shape and size of the RF remains constant at all correlation values. Only amplitude scaling is changed. Models of LGN cells in which RF surround regions are time delayed illustrate this effect (Cai et al. 1997). For a time-delayed surround, the center RF subregion decreases in size with correlation delay. Overall RF size, which includes the surround, increases with correlation delay. The combination of RF center size decrease with overall size increase can account for increased frequency (resolution) and increased overall size of monocular RFs of simple cells.

Population data for the LGN cell recordings are given in Fig. 3. The histograms here show the change in size of the RF center during the temporal windows indicated. The distribution has an approximately Gaussian shape, but it includes a subgroup of cells (from −20 to −16%/10 ms) that exhibits a relatively large rate of decrease in RF center size. The average rate of decrease for the entire sample is 7.6%/10 ms. This is significantly different from no change (*P* < 0.0001). It is a relatively large effect compared with that for cells in the visual cortex and is greater by ∼3–4%/10 ms (see following text).

Having considered the dynamic properties of RF size, we turn now to the question of temporal latency. Specifically, do large LGN RFs have shorter latencies? This would imply that coarse processing precedes that on a fine scale. To examine this, we plot optimal delay times versus RF center sizes in Fig. 4. The clear tendency here is for cells with relatively large RF

FIG. 4. LGN cells with smaller center sizes are more delayed than those with larger center sizes. At the optimal time slice, the center size parameter from the Gaussian fit is plotted against the optimal delay. There is a weak correlation supporting the notion that lower frequency [larger receptive fields (RFs)] information is processed before higher frequency (smaller RFs) information (robust regression, *P* = 0.029, slope = −0.011). There are 5 cells with optimal time delays >60 ms. These data are excluded because the extreme latency cells are not consistent with the observed correlation.

centers to have shorter latencies (correlation coefficient = −0.42). The regression fit has a slope of −0.011°/ms (*P* = 0.029). This finding is consistent with data from cortical cells, as shown in the following text.

Simple cells

An example of disparity-tuning dynamics of a simple cell is shown in Fig. 5. Binocular components are given in Fig. 5, *A–D.* Monocular functions are depicted in *E–L.* Each set of data points for temporal slices at optimal and ±25 ms before and after optimal, are fitted with Gabor functions (Fig. 5, *A–C*). Inspection of these functions shows that flanking subregions move closer to the central peak as correlation delay increases. The position of the central peak remains constant for all delay times. At higher correlation delays, the RF subregions decrease in size and move closer together. These changes are expressed in terms of the Gabor function fits in disparity frequency (resolution) and size (range) in Fig. 5*D.* Data points, presented for 5-ms differences from optimal time (see following text), show a linear decrease and increase for size and frequency, respectively.

Our method provides simultaneous measurements of monocular and binocular functions by selective cross-correlation of the spike train to different stimuli. We are therefore able to compare temporal RF dynamics with respect to frequency (resolution) and range (size) for monocular and binocular conditions. In this case, the expectation is that disparity-tuning dynamics may be predicted by monocular responses. Monocular RF profiles for left and right eyes are displayed in Fig. 5, *E–K,* for the same time slice as those for disparity tuning. Rates of change for frequency and size for left and right eyes, are depicted in *H* and *L,* respectively. Response patterns for left and right eyes are reasonably well matched. For both eyes, flanking, bright-excitatory subregions become stronger and

FIG. 3. Distribution of LGN center size changes. - - -, 0 change. The mean rate of change is –7.6%/10 ms, which is significantly <0 (*P* < 0.0001).
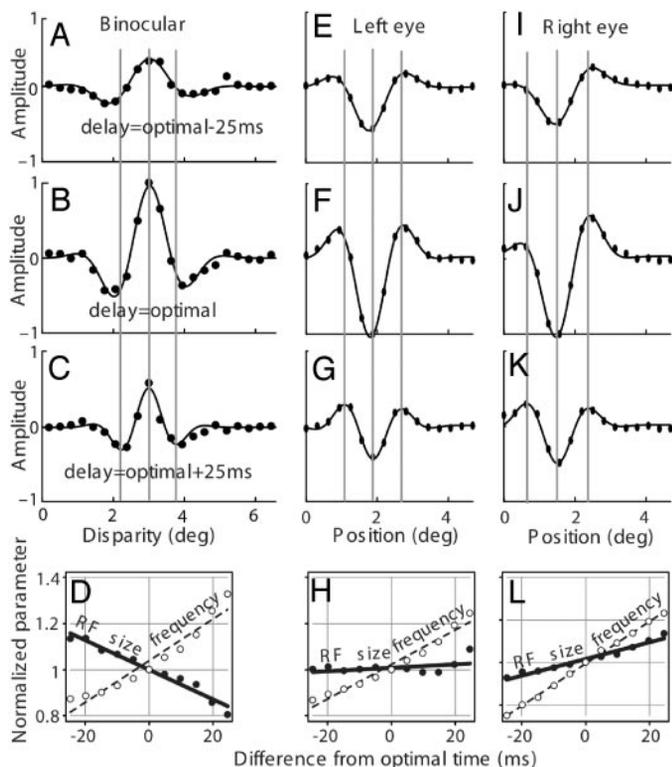
FIG. 5. Example of disparity-tuning dynamics and monocular RF dynamics for a simple cell. *A–D*: the suppressive flanking subregions move closer to the optimal disparity with greater correlation delay, which correlates with increasing frequency and decreasing size [frequency slope = 9.0 (%/10 ms); size slope = −6.4 (%/10 ms)]. *E–H*: the left eye 1-D receptive field is shown at the same correlation delays as the disparity-tuning curves. The flanking excitatory subregions become stronger with correlation delay. There is no change in RF size, but frequency increases [frequency slope = 7.4 (%/10 ms); size slope = 0.8 (%/10ms)]. *I–L*: the right eye 1-D receptive field is shown at the same correlation delays as the disparity-tuning curves. There is an increase in the number of subregions from 2 (*I*) to 3 (*K*) with increasing correlation delay. This effect is correlated to an increase in size and frequency [frequency slope = 9.9 (%/10 ms); size slope = 3.9 (%/10 ms)].

move closer to the central trough as correlation delay increases. Gabor function fits show corresponding increases in frequency (resolution) and size (range) although RF size for the left eye remains nearly constant. Because frequency and size factors increase, and the number of RF subregions is proportional to the product of these two factors, there is an increase in number of subregions as time delays progress from −25 to +25 ms. Comparison of binocular and monocular profiles reveals consistent patterns of frequency increases as time delays increase from −25 to +25 ms. However, the patterns for RF size changes are different for monocular and binocular cases. In particular, binocular RF size decreases with increasing time delays but monocular patterns are somewhat flat (left eye) or increase (right eye) with time. This difference in monocular and binocular patterns is considered in the following text.

Complex cells

The disparity-tuning dynamics of a complex cell are illustrated in Fig. 6. As in the previous example, tuning curves, fit with Gabor functions, are shown for temporal delays at the optimal time (*B*) and at −20 and +20 ms (*A* and *C*, respectively). Vertical lines mark the peak and troughs at the +20 ms

level (*C*). The vertical bars are extended into *A* and *B* to help show that the flanking RF subregions move closer to the central peak with increasing correlation delay. Subregions also become smaller and closer together. Note that for all 3 time delays, the position of the central peak remains constant. From
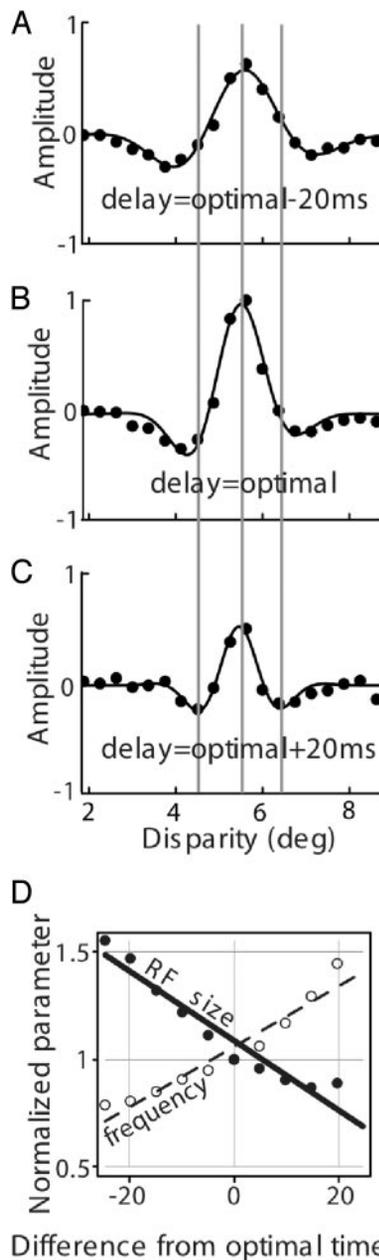
FIG. 6. Example complex cell demonstrating that disparity tuning becomes sharper with increasing correlation delay. *A–C*: disparity-tuning data (●) with normalized amplitudes are fit with Gabor functions (—) at 20 ms before the optimal delay (*A*), optimal delay (*B*), and 20 ms after the optimal delay (*C*). |, the optimal disparity is not changing, while the suppressive flanking subregions move closer to the optimal disparity at increasing correlation delay. *D*: the frequency (○, - - -) and RF envelope size parameters (●, —) of the Gabor fit to the disparity-tuning curves are shown at each correlation delay, normalized by the value of the parameter at optimal delay. Frequency and envelope size increase and decrease, respectively, in a generally monotonic manner. Both effects extend beyond the optimal delay. The change in these parameters as a function of time are fit with straight lines, and the slope is used to characterize the parameter's rate of change [frequency slope = 14.0 (%/10 ms); size slope = −16.2 (%/10 ms)].

the Gabor function fits, RF disparity frequency (resolution) and disparity range (size) may be estimated. The data (Fig. 6D) show that RF range (size) decreases approximately linearly with increasing time delay. A similar but opposite change is shown for frequency (resolution) that increases with increasing time delay. The results of Fig. 6 show clearly that optimal disparity, defined by the location of the main peak, does not change with correlation delay. This finding, combined with the increase in disparity frequency (resolution) and decrease in disparity range (size) constitutes a coarse-to-fine disparity process. This kind of mechanism was put forward in a model that proposed to account for how the visual system solves the retinal disparity correspondence problem (Marr and Poggio 1979). In the following text, we examine our cell population to see if the overall results are consistent with this notion.

Before doing this, we consider a methodological issue concerning temporal binning. Our stimulus duration is 40 ms, whereas the data shown so far are presented in 5-ms bins. We are therefore interpolating within the temporal RF, and this could cause an inaccuracy. To explore this, we used a bin duration equal to that of the stimulus. We use only two time delays (40 and 80 ms) because of the large bin size. Fewer Gabor function fits meet our criteria for analysis, so the sample size is reduced compared with that for 5-ms bins. Results of the comparison of 5- and 40-ms bin widths, for disparity frequency (resolution) and range (size) are shown in Fig. 7, A and B, respectively. For both parameters, a relatively close correlation is seen for results with the two bin widths [0.68 and 0.81 for frequency (resolution) and range (size), respectively]. This result demonstrates that the brief binning method is appropriate. This applies as long as the optimal delay is at the center of the analysis window and not at the end. In this way, there is an intrinsic control for any possible interpolation artifact. Interpolation between two valid time slices yields reliable results. An alternative method of analysis is to bin every 5 ms but only use data points that are 40 ms apart (Menz and Freeman 2003). This method produces similar results as we show in the following text. In a previous report, these issues have been clearly discussed (Ringach et al. 2003)

Distributions are shown in Fig. 8 for our population of complex cells for the main parameters of interest i.e., disparity frequency (resolution; A) and disparity range (size; B). Two binning times, as discussed in the preceding text, are presented in these distributions (5 and 40 ms in □ and ■, respectively). Distributions for the two binning methods are broadly similar and not significantly different ($P = 0.93$ and $P = 0.42$ Wilcoxon rank-sum test, for frequency and range distributions, respectively). This finding confirms the justification to use the shorter interpolated bin width, and subsequent data are presented in this form because it provides a larger sample size. For the entire distribution using the 5-ms binning method (Fig. 8A), average increase in disparity frequency is 4.5%/10 ms. This is a relatively modest but clear effect. Only 9 of 87 cells show decreases in frequency, and they are relatively small. The data for disparity range change, presented in Fig. 8B, also shows a clear trend, but it is weaker than that for frequency. The average decrease in disparity range is 3.2%/10 ms. In the case of range, 17 of 87 cells show a reverse effect, i.e., an increase rather than a decrease.
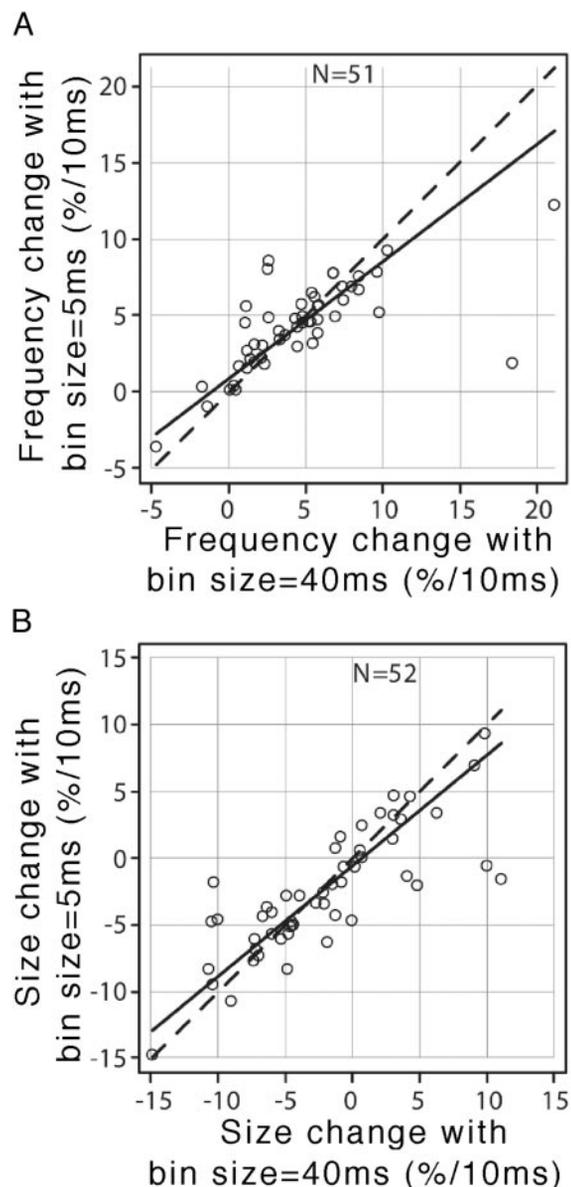

FIG. 7. A comparison of dynamics between different methods of binning for complex cells. The standard temporal binning (correlation delay intervals) is in 5-ms steps that produce a smooth profile in the temporal domain. However, this interpolation could produce artifactual dynamics. The duration of the stimulus is 40 ms, so 40-ms bins cannot have any artifact, but we are now limited to only 2 time slices. These figures compare frequency and size dynamics with the 2 types of binning methods. - - -, a 1-to-1 relationship; —, robust regressions. A: the rate of disparity frequency change for the interpolated method (5-ms bins) is plotted against the rate of disparity frequency change determined from 2 time slices that are 40 ms apart (i.e., noninterpolated method). There is a generally good correlation between the 2 methods (correlation coefficient = 0.68). The overall population trend does not change (Wilcoxon signed-rank, $P = 0.9551$). The robust regression does have a slope and offset that is significantly <1.0 and >0, respectively (slope = 0.77, offset = 0.92). The interpolated method underestimates at the high end (>5%/10 ms data tends to be slightly below the dashed line), while the overestimates tend to come at the low end (<5%/10ms data tend to be above the dashed line). So there is a slight tendency for the interpolated data to underestimate big dynamic changes and overestimate very small ones compared with the 2 slice noninterpolated method. B: the rate of disparity size change is compared using the 2 binning methods. The results are very similar to A in the preceding text (correlation coefficient = 0.81, Wilcoxon signed-rank, $P = 0.35$). The robust regression has a slope significantly <1 (slope = 0.83). There is slight tendency for the interpolated method to underestimate the rate of change compared with the noninterpolated method.
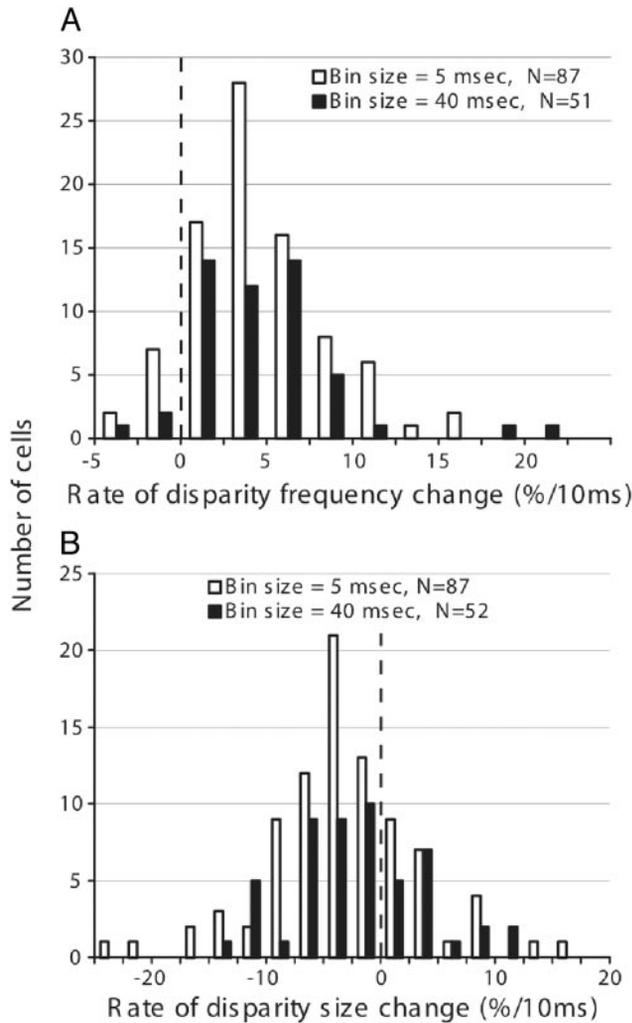
FIG. 8. Distributions of disparity frequency and disparity range (size) dynamics for all complex cells. The positions of the - - - (at 0) show that these distributions are shifted away from 0. □, the interpolated binning method (5 ms); ■, the noninterpolated method (40 ms). *A*: for the interpolated method, the mean disparity frequency increase is 4.5%/10 ms, and this is significantly >0 ($P < 0.0001$). The noninterpolated method has a distribution that is not significantly different from the interpolated method (Wilcoxon rank-sum $P = 0.93$). The mean frequency increase is 4.6%/10 ms and it is significantly >0 ($P < 0.0001$). *B*: for the interpolated method, the mean decrease in disparity size is 3.2%/10 ms, and this is significantly <0 ($P < 0.0001$). The noninterpolated method of binning has a distribution that is not significantly different from the interpolated method ($P = 0.42$, Wilcoxon rank-sum), Although the mean decrease is less (2.1%/10 ms), it is still significantly <0 ($P = 0.0046$).

Comparisons of dynamics of simple and complex cells

A primary question of this study, posed at the outset, is related to the notion of coarse-to-fine processing. One relevant parameter is the relation between optimal time delay and different stimulus parameters. Specifically, in a coarse-to-fine process, optimal time delays could be longer for high compared with low disparity frequencies. This means that coarse information would be processed first followed by that of fine detail. A comparison of optimal time delay and disparity frequency for our population of simple and complex cells is shown in Fig. 9, *A* and *B,* respectively. Although the distributions are relatively broad, there is a clear tendency, for both simple and complex cells, for optimal time delay to increase with disparity frequency. Robust

regression lines fit to the data have slopes of 0.0081 and 0.0085 cycle $\cdot$ $^{\circ -1}$ $\cdot$ ms^{-1} for simple (*A*) and complex (*B*) cells, respectively. These values are significantly different from zero ($P = 0.031$ and 0.0001, respectively). Clearly, neurons with higher preferred disparity frequencies are relatively more time delayed. The correlation is slightly stronger for complex compared with simple cells (correlation coefficients of 0.42 and 0.33, respectively). This coarse-to-fine dynamics could be accounted for by neurons that pool
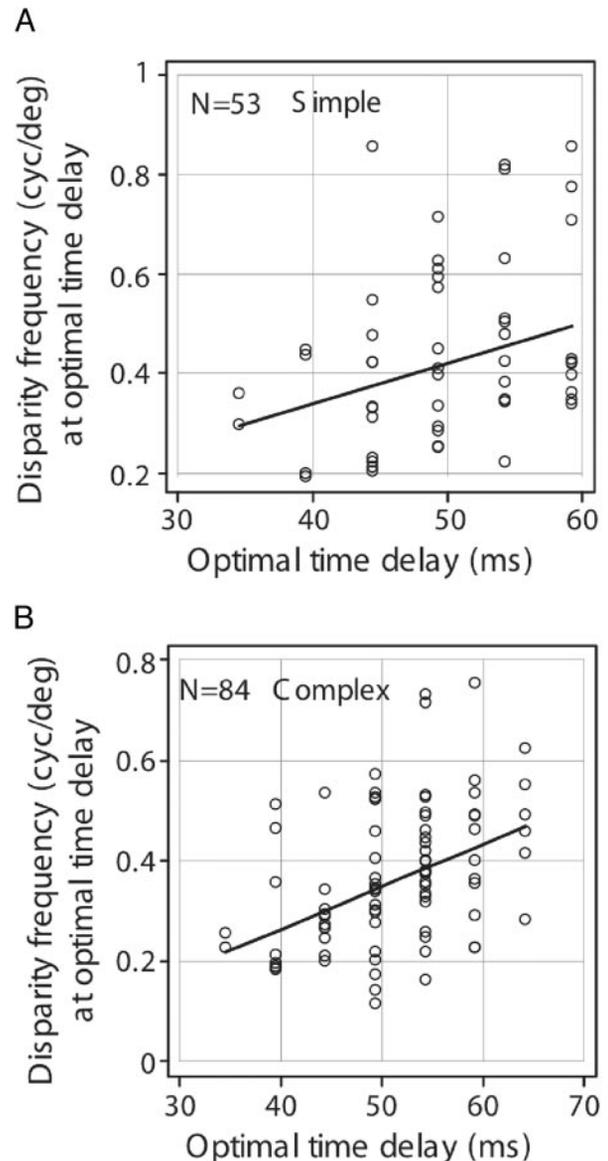

FIG. 9. Relationship between disparity frequency tuning and optimal time delay. *A*: this is the same relationship for simple cells, and 8 data points with optimal delays >60 ms were excluded for the same reason mentioned in the preceding text. The correlation between frequency and latency is weaker compared with complex cells ($P = 0.031$, with a slope = 0.0081 cycle $\cdot$ $^{\circ -1}$ $\cdot$ ms^{-1}). The magnitude of the regression slopes for complex and simple are nearly identical. *B*: the optimal disparity frequency measured at the optimal time delay is plotted against the optimal time delay for complex cells. Higher frequency cells are relatively more time delayed compared with lower frequency cells. The 5 data points with optimal delays >65 ms have been excluded because this frequency-latency correlation is not maintained at these very high time delays. The robust regression (—) has a slope (0.0085 cycle $\cdot$ $^{\circ -1}$ $\cdot$ ms^{-1}) that is significantly different from 0 ($P = 0.0001$).

input from other cells of slightly different disparity frequency content. The lower disparity frequency information is represented relatively earlier in the response. Our data for both complex and simple cells are consistent with this feed-forward mechanism for generating coarse-to-fine disparity tuning.

An overall comparison of the rates of changes in our population of cortical cells is given in the histograms of Fig. 10. Data are presented for rates of changes for disparity frequency (*A*) and disparity range (*B*) for monocular and binocular simple cell RFs and for complex cells. Monocular rates are the averages for left and right eye dynamics. We find no obvious bias for left and right eye responses. In general, distributions are quite similar for the three conditions illustrated. For disparity frequency (Fig. 10*A*), distributions are all skewed to positive rates of change, and there is not a significant difference between them ($P = 0.096$, ANOVA). Mean values of rates of disparity frequency increases are 4.2, 3.3, and 4.5%/10 ms for, respectively, monocular RFs of simple cells, disparity tuning of simple cells, and complex cells (*A*). Mean values for range changes with correlation delay are: a mean increase of 2.2%/10 ms, and mean decreases of 2.3 and 3.2%/10 ms for monocular
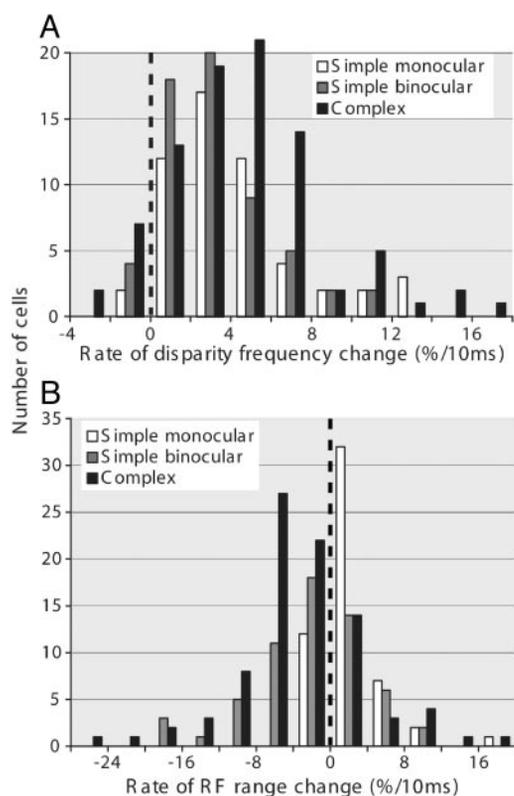

FIG. 10. A comparison of population dynamics between disparity-tuned simple and complex cells, and the monocular RFs of the disparity-tuned simple cells. - - -, 0 change. *A*: the change in frequency with correlation delay is not significantly different across all 3 types of receptive fields. ($P = 0.0959$, ANOVA). *B*: however, the change in RF range (size) is significantly different between RF types ($P < 0.0001$, ANOVA). The monocular RFs from the disparity-tuned simple cells generally increase in size with correlation delay, and this is opposite the direction of the disparity dynamics for the same cells. While the RF size decreases for disparity-tuned complex and simple cells are not significantly different ($P < 1.000$, Bonferroni), the monocular RF size changes from the disparity-tuned simple cells (□) are significantly different from both disparity-tuned complex (■, $P < 0.001$, Bonferroni) and disparity-tuned simple cells (▨, $P < 0.001$, Bonferroni).

RFs of simple cells, disparity tuning of simple cells, and complex cells, respectively (*B*). The decreases in binocular simple and complex cells are not significantly different from each other ($P < 1.000$, Bonferroni), but they are both different from the monocular RFs of simple cells ($P < 0.01$, $P < 0.001$, Bonferroni).

Having addressed the rates of changes of disparity frequency and range, we now want to know if these dynamics are correlated. If there are associated changes in these two variables, then there may be a common underlying mechanism. The relevant data for simple cell monocular RFs, simple cell binocular RFs and complex cell RFs are shown in Fig. 11, *A–C,* respectively. Robust regression line fits to the data show the relevant slopes that are not significantly different in the case of simple cell binocular RFs (*B*) and complex cells (*C*). In the case of monocular simple cells (*A*), the regression slope is opposite, i.e., relatively large frequency increases tend to be associated with large RF range increases. In all three cases, weak correlations apply: Pearson correlation coefficients of 0.12, -0.43, and -0.21, respectively, for monocular RFs of disparity-tuned simple cells, binocular disparity-tuned simple cells, and complex cells. We conclude that rates of changes in frequency (resolution) and range (size) are weakly coupled.

As noted in the preceding text, there is a discrepancy between the monocular and disparity-tuning dynamics from the same simple cells. Specifically, disparity frequency changes are similar in the two conditions, but range (size) changes occur in opposite directions. Our measurements provide simultaneous data on monocular RFs and binocular disparity tuning from simple cells that allow direct comparisons between the dynamics of these two conditions. Changes in disparity frequency and range (size) are shown, respectively, in Fig. 12, *A* and *B,* for monocular versus binocular conditions. One-to-one conditions designated (- - -, slopes of 1). For disparity frequency changes (*A*), monocular and binocular data are correlated (correlation coefficient = 0.44) and the slope is significantly different from 1.0 (robust regression, slope = 0.51). Monocular frequency changes tend to be slightly higher than those of disparity frequency (mean difference = 0.82, $P = 0.03$). Regarding range (size) changes (*B*), correlation between monocular and binocular values is relatively weaker (correlation coefficient = 0.28). Monocular range (size) changes are higher than those for binocular conditions (mean difference of 4.0, $P < 0.0001$). This difference is clear graphically as all data points for decreasing disparity range (size) to the left of zero fall above the dashed one-to-one line. Overall, monocular dynamics do not predict those of binocular disparity and this is more apparent in the range (size) domain compared with that of frequency. Specifically, our model of simple cells consists of left and right linear RFs that can be modeled as Gabor functions followed by linear summation and a static, half-squaring nonlinearity This model predicts that disparity dynamics should be a weighted sum of the monocular dynamics, but this is clearly not the case.

Finally, we consider differences in monocular dynamics for left and right eye RFs in a population of binocular simple cells. The relevant data are presented in Fig. 13. Of the four distributions, right and left eye spatial frequency changes (Fig. 13*C*) are the most correlated (correlation coefficient = 0.60). For these data, the slope (0.37, —) is significantly different from
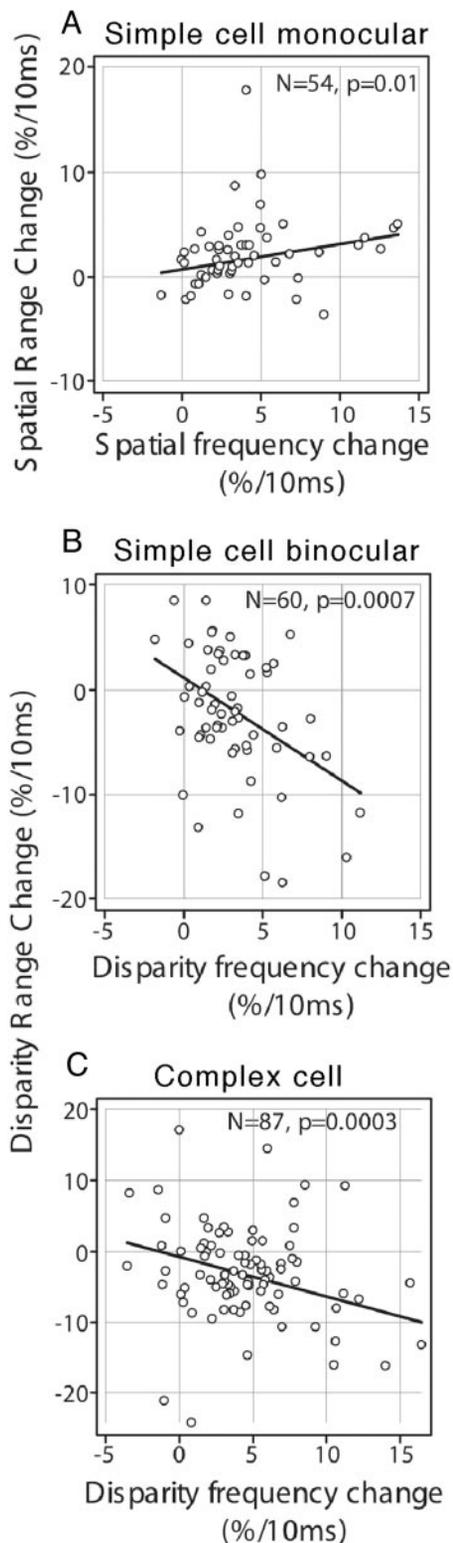
FIG. 11. The relationship between RF range (size) and disparity frequency dynamics. *A*: for monocular RFs of disparity-tuned simple cells, relatively large frequency increases tend to have large RF size increases. The frequency and size dynamics of disparity-tuned simple, *B,* and complex, *C,* cells exhibit similar weak correlations (robust regressions, the slopes are not significantly different). Cells with relatively large frequency increases tend to have large RF size decreases.
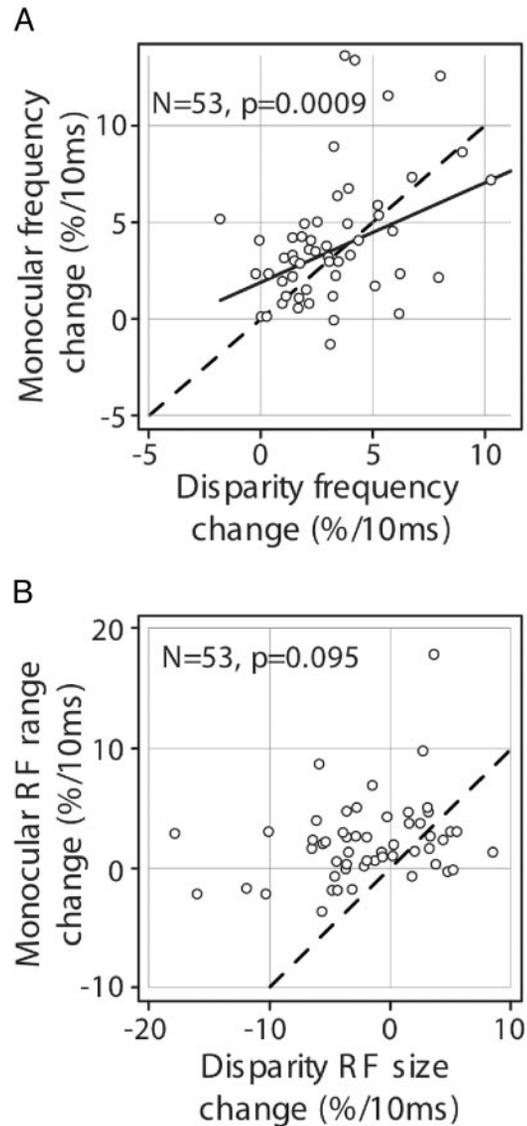

FIG. 12. The relationship between monocular and binocular dynamics of simple cells. - - -, a 1-to-1 relationship. *A*: monocular frequency changes are correlated to disparity frequency changes, but the slope is significantly different from 1.0 (robust regression, slope = 0.51). Furthermore, the monocular frequency changes tend to be slightly higher than the disparity frequency changes (mean difference = 0.82, $P = 0.03$). *B*: the monocular RF range (size) changes are not significantly correlated to the binocular RF size changes. The monocular size changes are much higher than the disparity size changes (mean difference: 4.0, $P < 0.0001$). The difference between monocular and disparity size changes is dramatic (all the data are above - - -) when the disparity size is decreasing (i.e., negative values). For the few cells with increasing disparity size, the data are distributed across the 1-to-1 - - -.

1.0 (- - -). For the other variable, RF range, (Fig. 13*B*) the distribution for left and right eye values is more scattered with a weak correlation (correlation coefficient = 0.25). As in the case of the previous analysis, these monocular comparisons show relative similarity in the frequency change domain and variability in RF range. The other comparisons provided in Fig. 13 are for spatial frequency and RF range changes for each eye. In this case, there is a weak correlation for the right eye (*A*; correlation coefficient = 0.25) and a weaker one for the left eye (*D*) (correlation coefficient = 0.16). In the latter case, there is not a significant regression slope.
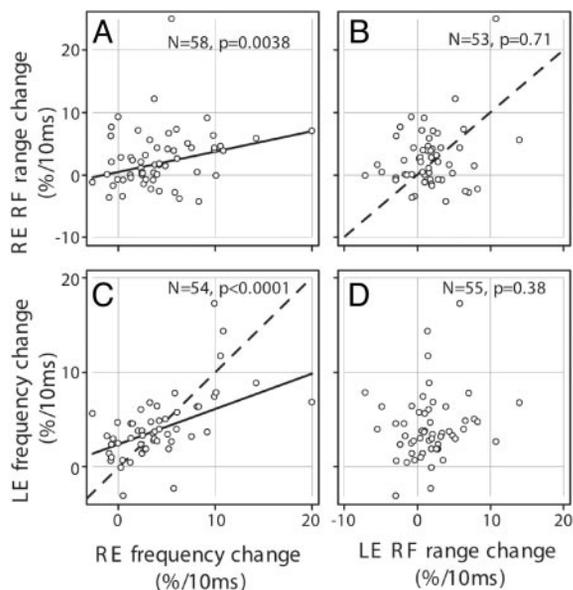
FIG. 13. A comparison of left and right eye RF dynamics. This matrix of scatterplots compares left and right eye RF range and frequency changes from binocular simple cells. *A*: range and frequency changes are correlated for right eye RFs (robust regressions slope = 0.33). *B*: left and right eye RF range changes are not correlated. *C*: left and right eye frequency changes are correlated. The slope is significantly different from 1.0 (slope = 0.37). *D*: range and frequency changes for the left eye RFs are not correlated according to the robust regression algorithm, which is supposed to minimize the influence of outliers.

DISCUSSION

We have examined temporal and spatial properties of neurons in the central visual pathway to explore processing characteristics for binocular disparity. The dynamics of LGN cells were examined to determine how RF center size changes during the course of the response. Even though these cells are essentially monocular, we assume they provide the input to simple cells in the visual cortex that we presume form the first stage of disparity encoding. We determined disparity frequency (resolution) and disparity range (size) of the monocular and the binocular RFs of simple cells that were disparity tuned. We also studied the temporal dynamics of these cells and of complex cells. The primary result is that there is a temporal progression of neural processing that begins with low disparity frequency coarse detail and ends with high-resolution information. This temporal sequence occurs while the optimal disparity remains constant. What is especially interesting about this observation is that it is consistent with a well-known theoretical proposal by which the stereoscopic correspondence problem may be solved by a coarse-to-fine process. (Marr and Poggio 1979). Note, however, that the coarse-to-fine mechanism that we have investigated covers a relatively narrow range. To eliminate false matches, this process requires a considerable narrowing of the range. It is possible that other mechanisms beyond V1 effectively expand the range of this function. In any case, our results leave open the possibility that this coarse-to-fine process is not restricted to the correspondence problem. It may apply to the entire stereoscopic process, and it also could underlie other visual functions. Visual perceptions could involve temporal sequences by which low spatial frequency information is processed first providing an initial coarse view which is then refined with time.

The mechanism for this process could begin at or prior to the

LGN. Our present findings show that LGN cells exhibit a decrease in center size with correlation delay. This result is associated with a time-delayed LGN RF surround. This effect could contribute to the increased Gabor function spatial frequency with time that is exhibited in monocular RFs of simple cells. In general, disparity dynamics of simple and complex cells are similar. This could be the result of pooled input from simple to complex cells. However, there is a discrepancy between monocular and binocular dynamics of disparity-sensitive simple cells. Disparity frequency dynamics are similar in the two conditions but range (size) temporal changes occur in opposite directions, i.e., monocular RFs increase in size while disparity range (size) decreases. This effect is considered below.

Tuning dynamics

Important theoretical concepts have been put forward to account for motion detection and related perceptual phenomena. Energy models may be most suitable for this purpose (Adelson and Bergen 1985; Watson and Ahumada 1985). These models have been modified to provide a theoretical basis for the processing of stereoscopic information (Anzai et al. 2001; DeAngelis et al. 1995; Freeman and Ohzawa 1990; Ohzawa et al. 1997), and experimental tests have been conducted to determine if predictions of the models are matched by the resulting data. Most of these tests are concerned with changes in spatial location of a subregion over time. In the work reported here, we are concerned with changes in tuning characteristics over a relatively limited time epoch (40 ms).

Previous studies have been conducted in which the dynamics of tuning characteristics of cortical cells have been examined. The most common characteristic is orientation tuning. Results of these studies are mixed. Orientation bandwidth has been reported to become narrower during the temporal course of the response of cortical cells (Best et al. 1989; Ringach et al. 1997; Volgushev et al. 1995). In other experimental work, orientation tuning characteristics are reported to be largely unchanged during the temporal course of the response (Celebrini et al. 1993; Mazer et al. 2002). The spatially changing positive effects were from anesthetized preparations and the time-constant findings were from awake behaving animals. It is not clear if this difference is relevant to the findings. The question of temporal dynamics has also been applied to spatial frequency tuning. In this case, it has been reported that tuning preference changes from lower to higher spatial frequencies over time (Bredfeldt and Ringach 2000). This latter finding is consistent with our current results. Finally, a recently reported study of spatial dynamics of RFs in visual cortex utilized relatively long-duration stimuli (300 ms). The reported finding is that RF subregion width decreased with greater delays in a reverse correlation procedure (Suder et al. 2002). Their data are more consistent with a feed–forward model from thalamic input than with one that includes intracortical feedback. A feed–forward mechanism is what we propose to account for the results of the current study.

In addition to the type of preparation question noted in the preceding text, there are other experimental differences in the approach to this problem. For the recent studies, a type of reverse correlation stimulus procedure has been used (DeAngelis et al. 1993a,b). Generally, this means that a spike train is

cross-correlated to a noise stimulus sequence. In the case of our present study, the analysis is based on a temporal bin size that is equal to the stimulus duration. Our results of coarse-to-fine tuning are clear and consistent across the population we have studied. The other relevant result of our study is that monocular and binocular RFs of simple cells change size in opposite directions with time. This would not occur if the basic result was due to an artifactual dynamic.

Mechanisms

The original qualitative description of the stages of processing in central visual pathways implied a feed-forward system. Information was thought to be processed in a hierarchical manner through first-, second-, and third-order cells (Hubel and Wiesel 1962). More recent work has emphasized the possibility of intracortical processes that could involve feedback as in recurrent excitation models. (Ben-Yishai et al. 1995; Douglas et al. 1995; Somers et al. 1995). In the feed-forward process, the tuning of a postsynaptic cell is determined by pooling of information from many presynaptic cells. In the recurrent excitation model, feed-forward connections provide weak initial tuning that is then refined by feedback connections from neighboring cells, which also contribute broad based inhibition.

Either of these two mechanisms could produce a coarse-to-fine process such as the one we have described in the current study. Feed-forward processes could involve pooling of input from cells that have coarse-to-fine dynamics and pooling of input from cells with different spatial frequency content in which low-frequency latencies are shorter than those for high frequencies. Of course, both mechanisms may be involved. Although most of our current data may be accounted for by both feed-forward and feedback processes, this does not apply to the discrepancy between the monocular and binocular RF dynamics of simple cells. The feed-forward model cannot explain the discrepancy between the monocular RF and disparity-tuning size dynamics of binocular simple cells. Mathematically, this difference could be described as an exponent on the nonlinearity that increases with time. The greater the exponent, the smaller the disparity size becomes, without altering monocular RF size. Biologically, it is reasonable to speculate that there is disparity-tuned feedback from either complex cells or multiple simple cells. This results in monocular tuning that does not entirely predict disparity tuning as in the case of complex cells.

How does a coarse-to-fine process work? Low spatial frequency filters can respond to a wide range of disparities but with poor resolution. High spatial frequency filters have fine resolution but can only respond accurately to a limited range of disparities. Prior to the processing of disparity information, the system must solve a correspondence problem, i.e., it must select the correct right-left match so that the appropriate depth plane is identified. To do this, frequency and range information must be taken into account. An important theoretical proposal was put forward to address this issue. Information across spatial frequency scale may be combined so that low-frequency information constrains the range and this is followed by high-frequency resolution (Marr and Poggio 1979). The order of disparity information processing thus follows a coarse-to-fine sequence. The data we provide here are consistent with this notion. Our neurophysiological findings suggest a pooling across spatial frequency scale with a temporal bias from low to high frequencies that causes a coarse-to-fine process.

The original coarse-to-fine stereoscopic processing theory was followed by some refinements and variations (e.g., Nishihara 1984; Nomura 1993; Qian and Zhu 1997). A number of behavioral studies have been conducted to explore the relationships between low and high spatial frequency processing to determine if the data are compatible with the theory. There is clear psychophysical evidence that low spatial frequency information constrains processing on a fine scale (Rohaly and Wilson 1993; Wilson et al. 1991). On the other hand, some studies also suggest a reverse process by which high spatial frequency information is used to disambiguate that at low frequencies (Mallot et al. 1996; Smallman 1995; Smallman and MacLeod 1997). These processes could both occur by pooling across spatial frequency and averaging the result. In theory, this type of process can produce an unambiguous representation of disparity (Fleet et al. 1996). Disparity averaging across spatial scale has also been demonstrated psychophysically (Rohaly and Wilson 1994). Another approach to the idea of coarse-to-fine processing is to examine the temporal order of spatial frequency processing. In this case results show that low-frequency information is processed more rapidly than that of high values, i.e., there is a temporal coarse-to-fine mechanism (Glennerster 1996; Watt 1987). An additional relevant study shows that there is a transient stereopsis process that is temporally fast and consists of low spatial frequency information (Schor et al. 1998). This again is consistent with a coarse-to-fine process. It is important to point out that this type of mechanism may apply to other visual functions such as object recognition (Parker et al. 1997; Watt 1987). Considered together, the theoretical, behavioral, and neurophysiological studies point strongly to a processing system that begins with an approximation and ends with a fine-tuned percept.

REFERENCES

Adelson EH and Bergen JR. Spatiotemporal energy models for the perception of motion. *J Opt Soc Am A* 2: 284–299, 1985.

Anderson CH and Van Essen DC. Shifter circuits: a computational strategy for dynamic aspects of visual processing. *Proc Natl Acad Sci USA* 84: 6297–6301, 1987.

Anzai A, Ohzawa I, and Freeman RD. Neural mechanisms for encoding binocular disparity: receptive field position versus phase. *J Neurophysiol* 82: 874–890, 1999a.

Anzai A, Ohzawa I, and Freeman RD. Neural mechanisms for processing binocular information. I. Simple cells. *J Neurophysiol* 82: 891–908, 1999b.

Anzai A, Ohzawa I, and Freeman RD. Joint-encoding of motion and depth by visual cortical neurons: neural basis of the Pulfrich effect. *Nat Neurosci* 4: 513–518, 2001.

Ben-Yishai R, Bar-Or RL, and Sompolinsky H. Theory of orientation tuning in visual cortex. *Proc Natl Acad Sci USA* 92: 3844–3848, 1995.

Best J, Mallot H, Kruger K, and Dinse HRO. Dynamics of visual information processing in cortical systems. In: *Neural Networks: From Models to Applications*, edited by Personnaz L and Dreyfus G. Paris, France: IDSET, 1989, p. 107–116.

Bishop PO, Kozak W, and Vakkur GJ. Some quantitative aspects of the cat's eye: axis and plane of references, visual field coordinates and optics. *J Physiol* 163: 466–502, 1962.

Bredfeldt CE and Ringach DL. Dynamics of V1 neurons in the fourier domain. *Soc Neurosci Abstr* 26: 2090, 2000.

Bredfeldt CE and Ringach DL. Dynamics of spatial frequency tuning in macaque V1. *J Neurosci* 22: 1976–1984, 2002.

Cai D, DeAngelis GC, and Freeman RD. Spatiotemporal receptive field organization in the lateral geniculate nucleus of cats and kittens. *J Neurophysiol* 78: 1045–1061, 1997.

Celebrini S, Thorpe S, Trotter Y, and Imbert M. Dynamics of orientation coding in area V1 of the awake primate. *Vis Neurosci* 10: 811–825, 1993.

DeAngelis GC, Ohzawa I, and Freeman RD. Neuronal mechanisms underlying stereopsis: how do simple cells in the visual cortex encode binocular disparity? *Perception* 24: 3–31, 1995.

DeAngelis GC, Ohzawa I, and Freeman RD. Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. I. General characteristics and postnatal development. *J Neurophysiol* 69: 1091–1117, 1993a.

DeAngelis GC, Ohzawa I, and Freeman RD. Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. II. Linearity of temporal and spatial summation. *J Neurophysiol* 69: 1118–1135, 1993b.

Douglas RJ, Koch C, Mahowald M, Martin KA, and Suarez HH. Recurrent excitation in neocortical circuits. *Science* 269: 981–985, 1995.

Fleet DJ, Wagner H, and Heeger DJ. Neural encoding of binocular disparity: energy models, position shifts and phase shifts. *Vision Res* 36: 1839–1857, 1996.

Freeman RD and Ohzawa I. On the neurophysiological organization of binocular vision. *Vision Res* 30: 1661–1676, 1990.

Glennerster A. The time course of 2-D shape discrimination in random dot stereograms. *Vision Res* 36: 1955–1968, 1996.

Hubel DH and Wiesel TN. Receptive fields, binocular interaction and functional architecture of the cat's visual cortex. *J Physiol* 160: 106–154, 1962.

Jones JP and Palmer LA. The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *J Neurophysiol* 58: 1187–1211, 1987.

Mallot HA, Gillner S, and Arndt PA. Is correspondence search in human stereo vision a coarse-to-fine process? *Biol Cybern* 74: 95–106, 1996.

Marr D and Poggio T. A computational theory of human stereo vision. *Proc R Soc Lond B Biol Sci* 204: 301–328, 1979.

Mazer JA, Vinje WE, McDermott J, Schiller PH, and Gallant JL. Spatial frequency and orientation tuning dynamics in area V1. *Proc Natl Acad Sci USA* 99: 1645–1650, 2002.

Menz MD and Freeman RD. Stereoscopic depth processing in the visual cortex: a coarse-to-fine mechanism. *Nat Neurosci* 6: 59–65, 2003.

Nishihara HK. Practical real-time imaging stereo matcher. *Opt Eng* 23: 536–545, 1984.

Nishihara M and Kimura F. Differential influences of adrenoceptor agonists on the excitability of arcuate and medial preoptic neurons. *Nippon Juigaku Zasshi* 49: 1154–1156, 1987.

Nomura M. A model for neural representation of binocular disparity in striate cortex: distributed representation and veto mechanism. *Biol Cybern* 69: 165–171, 1993.

Ohzawa I, DeAngelis GC, and Freeman RD. Stereoscopic depth discrimination in the visual cortex: neurons ideally suited as disparity detectors. *Science* 249: 1037–1041, 1990.

Ohzawa I, DeAngelis GC, and Freeman RD. Encoding of binocular disparity by complex cells in the cat's visual cortex. *J Neurophysiol* 77: 2879–2909, 1997.

Parker DM, Lishman JR, and Hughes J. Evidence for the view that temporospatial integration in vision is temporally anisotropic. *Perception* 26: 1169–1180, 1997.

Press WH, Teukolsky SA, Vetterling WT, and Flannery BP. *Numerical recipes in C.* Cambridge, UK: Cambridge Univ. Press, 1992.

Qian N and Andersen RA. Transparent motion perception as detection of unbalanced motion signals. II. Physiology. *J Neurosci* 14: 7367–7380, 1994.

Qian N, Andersen RA, and Adelson EH. Transparent motion perception as detection of unbalanced motion signals. I. Psychophysics. *J Neurosci* 14: 7357–7366, 1994.

Qian N and Zhu Y. Physiological computation of binocular disparity. *Vision Res* 37: 1811–1827, 1997.

Quam LH. *Readings in Computer Vision.* Los Altos, CA: Kauffman, 1987.

Ringach DL, Hawken MJ, and Shapley R. Dynamics of orientation tuning in macaque primary visual cortex. *Nature* 387: 281–284, 1997.

Ringach DL, Hawken MJ, and Shapley R. Dynamics of orientation tuning in macaque V1: the role of global and tuned suppression. *J Neurophysiol* 90: 342–352, 2003.

Rohaly AM and Wilson HR. Disparity averaging across spatial scales. *Vision Res* 34: 1315–1325, 1994.

Rohaly AM and Wilson HR. Nature of coarse-to-fine constraints on binocular fusion. *J Opt Soc Am A* 10: 2433–2441, 1993.

Schor CM, Edwards M, and Pope DR. Spatial-frequency and contrast tuning of the transient-stereopsis system. *Vision Res* 38: 3057–3068, 1998.

Skottun BC, De Valois RL, Grosof DH, Movshon JA, Albrecht DG, and Bonds AB. Classifying simple and complex cells on the basis of response modulation. *Vision Res* 31: 1079–1086, 1991.

Smallman HS. Fine-to-coarse scale disambiguation in stereopsis. *Vision Res* 35: 1047–1060, 1995.

Smallman HS and MacLeod DI. Spatial scale interactions in stereo sensitivity and the neural representation of binocular disparity. *Perception* 26: 977–994, 1997.

Somers DC, Nelson SB, and Sur M. An emergent model of orientation selectivity in cat visual cortical simple cells. *J Neurosci* 15: 5448–5465, 1995.

Suder K, Funke K, Zhao Y, Kerscher N, Wennekers T, and Worgotter F. Spatial dynamics of receptive fields in cat primary visual cortex related to the temporal structure of thalamocortical feedforward activity. Experiments and models. *Exp Brain Res* 144: 430–444, 2002.

Sutter EE. The fast m-transform: a fast computation of cross-correlations with binary m-sequences. *SIAM J Comput* 20: 686–694, 1991.

Theunissen FE, David SV, Singh NC, Hsu A, Vinje WE, and Gallant JL. Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. *Network: Comput Neural Syst* 12: 289–316, 2001.

Volgushev M, Vidyasagar TR, and Pei X. Dynamics of the orientation tuning of postsynaptic potentials in the cat visual cortex. *Vis Neurosci* 12: 621–628, 1995.

Watson AB and Ahumada AJ Jr. Model of human visual-motion sensing. *J Opt Soc Am A* 2: 322–341, 1985.

Watt RJ. Scanning from coarse to fine spatial scales in the human visual system after the onset of a stimulus. *J Opt Soc Am A* 4: 2006–2021, 1987.

Wilson HR, Blake R, and Halpern DL. Coarse spatial scales constrain the range of binocular fusion on fine scales. *J Opt Soc Am A* 8: 229–236, 1991.