

Encoding of Binocular Disparity by Complex Cells in the Cat's Visual Cortex

IZUMI OHZAWA, GREGORY C. DEANGELIS, AND RALPH D. FREEMAN

Group in Vision Science, School of Optometry, University of California, Berkeley, California 94720-2020

Ohzawa, Izumi, Gregory C. DeAngelis, and Ralph D. Freeman.

Encoding of binocular disparity by complex cells in the cat's visual cortex. *J. Neurophysiol.* 77: 2879–2909, 1997. To examine the roles that complex cells play in stereopsis, we have recorded extracellularly from isolated single neurons in the striate cortex of anesthetized paralyzed cats. We measured binocular responses of complex cells using a comprehensive stimulus set that encompasses all possible combinations of positions over the receptive fields for the two eyes. For a given position combination, stimulus contrast could be the same for the two eyes (2 bright or 2 dark bars) or opposite (1 bright and 1 dark). These measurements provide a binocular receptive field (RF) profile that completely characterizes complex cell responses in a joint domain of left and right stimulus positions. Complex cells typically exhibit a strong selectivity for binocular disparity, but are only broadly selective for stimulus position. For most cells, selectivity for disparity is more than twice as narrow as that for position. These characteristics are highly desirable if we assume that a disparity sensor should exhibit position invariance while encoding small changes in stimulus depth. Complex cells have nearly identical binocular RFs for bright and dark stimuli as long as the sign of stimulus contrast is the same for the two eyes. When stimulus contrast is opposite, the binocular RF also is inverted such that excitatory subregions become suppressive. We have developed a disparity energy model that accounts for the behavior of disparity-sensitive complex cells. This is a hierarchical model that incorporates specific constraints on the selection of simple cells from which a complex cell receives input. Experimental data are used to examine quantitatively predictions of the model. Responses of complex cells generally agree well with predictions of the disparity energy model. However, various types of deviations from the predictions also are found, including a highly elongated excitatory region beyond that supported by a single energy mechanism. Complex cells in the visual cortex appear to provide a next level of abstraction in encoding information for stereopsis based on the activity of a group of simple-type subunits. In addition to exhibiting narrow disparity tuning and position invariance, these cells seem to provide a partial solution to the stereo correspondence problem that arises in complex natural scenes. Based on their binocular response properties, these cells provide a substantial reduction in the complexity of the correspondence problem.

INTRODUCTION

One of the most remarkable features of the visual system is the ability to see the world in three-dimensional depth. The visual system reconstructs depth from the pair of two-dimensional images projected on the retinas of the two eyes. These two images are very similar, but they contain small variations in the position of corresponding features in the visual scene, because the two eyes see the world from slightly different view points. This positional variation is called binocular disparity. Although there are other means

for estimating depth, stereopsis, the process of recovering depth information from binocular disparity, is usually the most robust and accurate for near distances (Howard and Rogers 1995; Pierce and Benton 1975). It has been demonstrated that binocular disparity alone can give rise to a vivid sensation of depth without the presence of any other depth cues (Julesz 1960, 1971; Wheatstone 1838).

The neural analysis of visual information for stereopsis is thought to begin in the primary visual cortex because it is the first stage along the visual pathway where neurons may be activated by stimulation of either eye and because extensive binocular interactions occur between stimuli presented to the two eyes simultaneously (Barlow et al. 1967; Ferster 1981; Hubel and Wiesel 1962, 1968; LeVay and Voigt 1988; Nikara et al. 1968; Ohzawa and Freeman 1986a,b; Ohzawa et al. 1990; Pettigrew et al. 1968; Poggio and Fischer 1977; von der Heydt et al. 1978). Although there are numerous studies that present descriptions of how neurons respond to binocular stimuli, little is known as to the specific neural circuitry that endows these neurons with the ability to respond to stereoscopic stimuli. We do not yet know, for example, the roles that simple and complex cells play with respect to stereopsis. Although both types of neurons clearly are tuned for binocular disparities (Ferster 1981; Joshua and Bishop 1970; Pettigrew et al. 1968), only complex cells appear to respond selectively to dynamic random dot stereograms (DRDS), which are defined by binocular disparity alone (Poggio et al. 1985, 1988; Poggio and Poggio 1984; Poggio 1995). These findings suggest that complex cells may perform more advanced and specialized processing of binocular information for stereopsis than simple cells.

Because the overall size of a complex cell receptive field (RF) is much larger than its optimal width for a bar-shaped stimulus (Emerson et al. 1987; Gaska et al. 1994; Movshon et al. 1978a), it is expected that multiple image features in the visual scene (with sizes optimally excitatory to the cell) will fall into the RF of each complex cell (Fig. 1A). Unlike simple cells, which have multiple discrete flanks and appear capable of signaling the presence of multiple features within the RF (Fig. 1B) by a linear transform (Gabor 1946; Geisler and Hamilton 1986; Ohzawa et al. 1996; Robson 1983; Watson 1991), complex cells seem to face a more difficult *binocular correspondence problem*, i.e., identification of corresponding image features in left and right images. This is one of the key problems in stereopsis (Julesz 1968, 1971; Marr and Poggio 1976). As illustrated in Fig. 1C (Julesz 1968, 1971), the correspondence problem arises because there is inherent ambiguity in attempts to match corresponding features in left and right images. Without an appropriate filtering

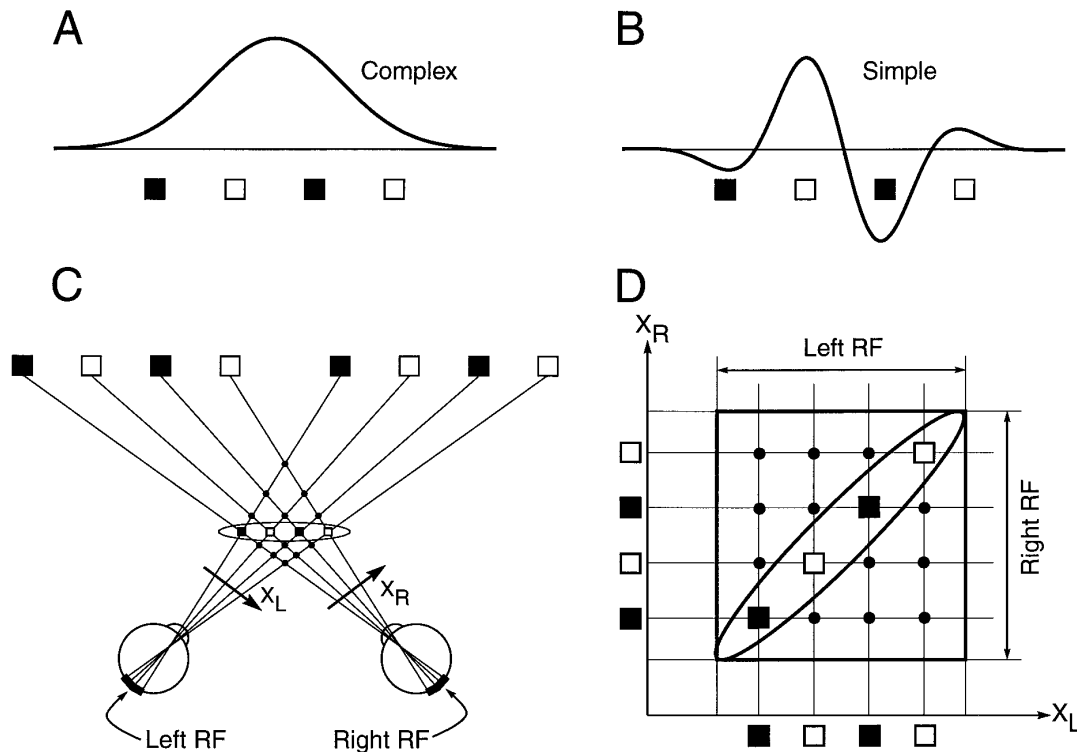


FIG. 1. Binocular correspondence problem is illustrated in relation to 2 major receptive field (RF) types: simple and complex. *A*: complex cells generally have large RFs that may contain several targets (bars) as illustrated (\square and \blacksquare). *B*: simple cells, on the other hand, have discrete RF subregions. *C*: when multiple targets are present as in *A*, there are large number of possible binocular matches (all intersections of rays) including false matches (\bullet) and correct matches (enclosed in a horizontal ellipse), (redrawn after Julesz 1971). Number of possible matches grows with square of number of targets in each image. Any stereo vision system must be able to solve the binocular correspondence problem, i.e., that of finding correct matches for left and right targets. If a complex cell responds in a nondiscriminatory manner to any conjunction of left and right targets within its RFs, it will not be able to tell the difference between correct and false matches. *D*: configuration of *C* is shown again in a Cartesian coordinate system.

mechanism, *false targets* (dots in Fig. 1, *C* and *D*) may elicit as much excitation as correctly matched targets (open and filled squares in Fig. 1, *C* and *D*). An examination of Fig. 1*C* indicates that one of the appropriate filtering operations may be the selection of the matches contained within the horizontally elongated ellipse. When plotted in a familiar Cartesian coordinate system (X_L , X_R), the desired region of sensitivity for such a filter is an elongated diagonal region as shown in Fig. 1*D*. Such a sensitivity map may be described as the *binocular receptive field* of the cell. We examine quantitatively if complex cells possess binocular RF profiles similar to that shown in Fig. 1*D*. It is also of interest to know the effects of stimuli falling outside the elongated diagonal area. They could cause suppression or have no effect. We also examine the responses of models in the same (X_L , X_R) domain, and compare predicted responses with the binocular RFs obtained from cells.

In addition to studying response properties of complex cells as outlined above, we wish to devise a physiologically realistic model for the role of these cells in binocular processing. Such a model must be consistent with the physiological data we obtain, as well as with findings from previous studies. We know that the RFs of complex cells appear generally broad in spatial extent, and nonspecific with respect to the sign of contrast (bright or dark) of a bar or edge stimulus (DeAngelis et al. 1995b; Hubel and Wiesel 1962;

Movshon et al. 1978a; Ohzawa et al. 1990). However, they possess similar spatial frequency tuning properties to those of simple cells and only slightly broader orientation tuning characteristics (Gizzi et al. 1990; DeValois et al. 1982; Movshon et al. 1978b). When studied with binocular stimuli, their disparity tuning is often narrower than the overall RF size predicts (Joshua and Bishop 1970; Pettigrew et al. 1968). This high degree of selectivity is thought to originate from multiple underlying RF subunits (Gaska et al. 1987, 1994; Movshon et al. 1978a; Ohzawa and Freeman 1986b; Spitzer and Hochstein 1985; Szulborski and Palmer 1990). Properties of these RF subunits appear to be similar in every respect to those of simple cells: 1) Within a monocular RF of a complex cell, spatial antagonism between neighboring subregions may be demonstrated by studying interactions between two stimuli that are presented at a variety of spatial separations (Movshon et al. 1978a; Szulborski and Palmer 1990). 2) These two-stimulus interaction profiles accurately predict the orientation and spatial frequency tuning of the complex cell (Gaska et al. 1994; Movshon et al. 1978a; Szulborski and Palmer 1990). 3) RF subunits of complex cells combine input from the two eyes in a linear manner (Ohzawa and Freeman 1986b). Furthermore, electrical stimulation of LGN afferents evokes mostly polysynaptic excitation in pyramidal neurons, especially in layers 2+3 (Douglas and Martin 1991). These findings strongly suggest a hierar-

chical model of a complex cell that consists of a simple cell subunit stage and an output stage at which multiple subunits are combined.

On the basis of physiological findings outlined above, a number of monocular hierarchical models of complex cells has been proposed (Adelson and Bergen 1985; Emerson et al. 1992; Gaska et al. 1994; Pollen et al. 1989). In this study, we develop a model that is suitable for the binocular case (Ohzawa et al. 1990) and derive a theoretical framework in which physiological results may be compared quantitatively with model predictions. For this purpose, we obtain a complete characterization of response properties of complex cells in the joint space-disparity-time domain and then compare these properties with predictions of specific models (Fleet et al. 1996, 1997; Ohzawa et al. 1990; Qian 1994, 1997).

METHODS

Surgical methods, experimental apparatus, and neurophysiological recording procedures have been described in detail elsewhere (DeAngelis et al. 1993a; Ohzawa et al. 1996). Brief descriptions and procedures not described previously are presented here.

Surgical procedure

Adult cats (2–4 kg) were prepared for electrophysiological recording as follows. First, a subcutaneous injection was given of Atropine sulfate (0.2 mg/kg) and Acepromazine (1 mg/kg). Anesthesia was induced and maintained during surgery with halothane (2.5–3% in oxygen). Electrocardiogram (ECG) electrodes and a rectal temperature probe were installed. ECG and core temperature were monitored using a PC-based physiological monitoring system (Ghose et al. 1995), which logs heart rate and temperature automatically every 5 min. Catheters were inserted into femoral veins on two limbs for infusion of drugs and fluids. A glass tracheal cannula was inserted immediately after tracheostomy. A stereotaxic apparatus was used to position the animal's head securely. Lidocaine ointment (5%) was used at pressure points. The skull was exposed, and two small machine screws were inserted for use as electroencephalogram (EEG) electrodes. Then, a craniotomy was performed to access the central representation of the visual field in the striate cortex (Horsley-Clark P4 L2.5). The dura was removed carefully to allow insertion of microelectrodes. After this point, anesthesia was administered by intravenous injection of sodium thiamylal (Surital). Then, paralysis was induced with an initial dose of gallamine triethiodide (Flaxedil, 7–10 mg/kg), and the animal was placed under artificial respiration. Anesthesia for the rest of the recording session was maintained by a combination of nitrous oxide (70% mixed with oxygen) and Surital (1 mg · kg⁻¹ · hr⁻¹). Paralysis was maintained by continuous infusion of Flaxedil (10 mg · kg⁻¹ · hr⁻¹) in lactated Ringer solution containing 5% dextrose. To maintain a proper level of respiration, a CO₂ sensor (Hewlett-Packard 47210A) was used. For the remainder of the experiment, four physiological parameters: heart rate, temperature, end-tidal CO₂ level, and EEG amplitude, were displayed continuously and logged by the monitoring system (Ghose et al. 1995). The system provides voice warnings if any of these parameters exceed preset limits. Pupils were dilated with 1% atropine sulfate solution, and nictitating membranes were retracted with 5% phenylephrine HCl. Contact lenses of appropriate power with 4 mm artificial pupils were placed over the corneas. Locations of optic disks and the area centrales were mapped onto a tangent screen using a reversible ophthalmoscope.

Stimuli and data acquisition

A tangent screen with a back-projected bar stimulus was used for initial exploration of RFs. The position and orientation of the bar were controlled by a joystick to facilitate manual exploration. Visual stimuli for quantitative measurements were generated on a matched pair of cathode ray tube displays and presented dichoptically through half-silvered front-surface mirrors angled at 45° in front of the animal's eyes. The mean luminance of the displays was 45 cd · m⁻² by direct viewing and 17 cd · m⁻² as viewed through the half silvered mirrors. The screens were placed at 57 cm from the cat's eyes, at which distance they subtended 28 × 22°.

The displays were driven by a PC-based dichoptic visual stimulator with two graphics adapters (Imagraph) and had a spatial resolution of 1,024 × 804 pixels. The two displays were refreshed at a frame rate of 76 Hz. The timing of video frames for the two displays was synchronized by hardware so that the dichoptic stimuli might be delivered without onset asynchrony. The stimulator generated sync pulses that indicated stimulus onset and timing of temporal modulation. These pulses were recorded by the data acquisition system along with the spike data.

Conventional amplifiers, oscilloscopes, and audio speakers were used to monitor raw signals from the microelectrodes. Spike data and stimulus sync pulses were recorded by custom-built data acquisition systems. A separate computer was used to control experiments and to perform preliminary real-time analysis of incoming data. The PC-based visual stimulator was controlled via a serial port, and the acquired spike and stimulus sync data were received from the data acquisition systems via high-speed interfaces. Sufficient information to reconstruct each trial, as well as times of occurrence of all spikes and sync pulses, were saved to a file to allow flexible and complete reanalysis of data. The data were recorded with 1 ms (previous system) or 40-μs (current system) resolution.

Recording procedures

Tungsten-in-glass microelectrodes (Levick 1972) were used for extracellular recording from neurons in the striate cortex. To increase the chance of encountering cells, two electrodes were mounted in a protective guide tube. They were driven in parallel with a single microelectrode drive (Inchworm, Burleigh). To minimize tissue damage, the two electrodes were not glued together to allow cortical tissue to pass between them. After confirming under a microscope that the electrodes do not penetrate blood vessels on the cortical surface, a small amount of agar in warm Ringer solution was applied to stabilize the cortex. Then molten wax was applied over the agar and the surrounding cranial bone to form a sealed chamber. This provided additional stability and protected the agar from drying.

After isolation of spike waveforms, the position and approximate preferred orientation of the RFs (for each eye) were recorded using a bar stimulus projected on the tangent screen. Then, an interactive search program (DeAngelis et al. 1993a) was used to determine optimal parameters with a small circular patch of drifting sinusoidal grating. Initial estimates of the optimal spatial frequency, orientation, and the center location of RFs were obtained. These values were refined by subsequent quantitative measurements under computer control. As a routine procedure, orientation tuning, and direction selectivity were measured first for each eye. Then, spatial and temporal frequency tuning curves were obtained. Each curve was defined by 7–11 points, and a cubic spline procedure (Press et al. 1992) was used to locate the peaks of the tuning curves.

Binocular RF measurement and analysis

A reverse correlation procedure (DeBoer and Kuyper 1968; Eggermont et al. 1983; Jones and Palmer 1987a; Sutter 1974, 1975)

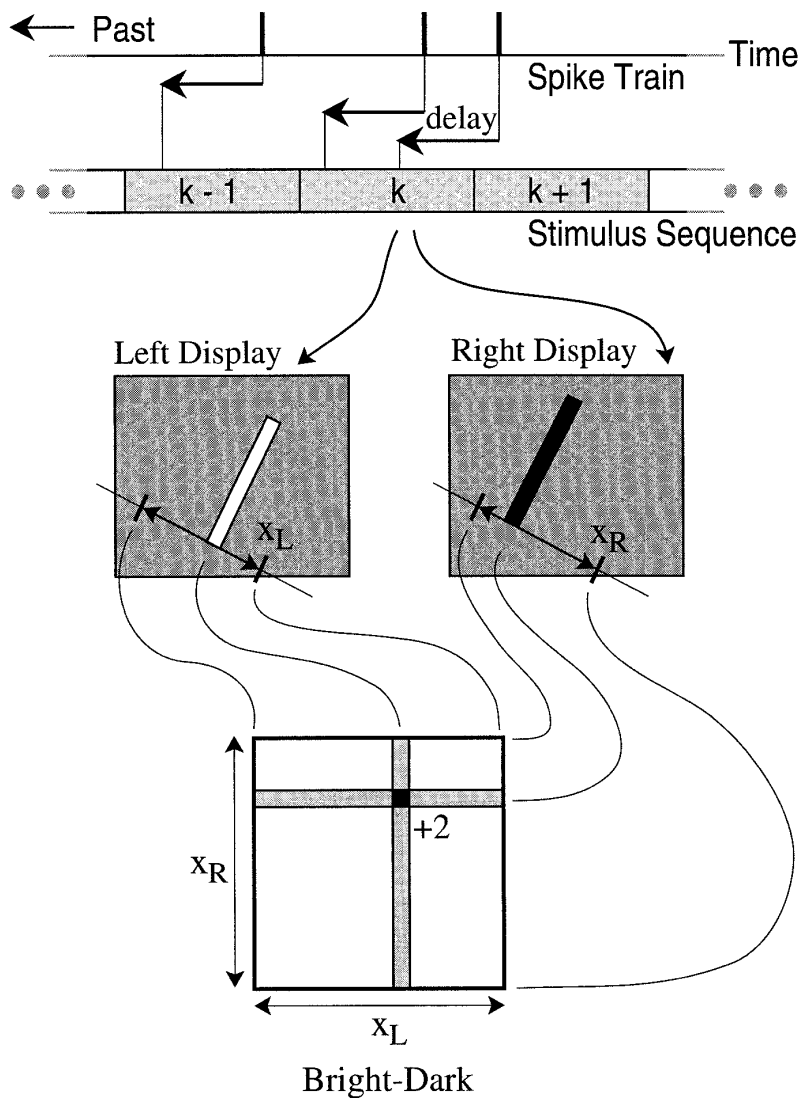


FIG. 2. A method of obtaining detailed binocular RF maps is illustrated. An exhaustive set of dichoptic stimuli is presented in which all possible combinations of left and right stimulus positions are included for each left-right permutation of contrast sign: dark-dark, bright-bright, dark-bright, and bright-dark. Stimuli in the set are presented in random order, 1 by 1, at a rate of 20–25 stimuli/s. Stimulus presentations are repeated after reshuffling the sequence. Resulting spike trains are cross-correlated with the stimulus sequence using a procedure known as reverse correlation (see text).

was used to study binocular RFs. Details of our monocular reverse correlation method have been described previously (DeAngelis et al. 1993a, 1995a; Freeman and Ohzawa 1990; Ohzawa et al. 1996). A binocular version of the reverse correlation method was similar to the monocular one except that each stimulus consists of a pair of bars presented dichoptically, as illustrated schematically in Fig. 2.

A binocular RF profile is defined in a joint two-dimensional domain (X_L , X_R) that includes the conjunction of left and right RFs (see Fig. 1D). Bright and dark bars of optimal orientations were presented in randomized order, at all possible combinations of left and right stimulus positions. Typically, 20 stimulus locations were used for each eye. This defined a 20×20 point stimulus grid in the (X_L , X_R) domain. Because each eye may be shown either a bright or dark bar, there were four permutations of bright and dark stimuli at each grid point. Therefore each binocular RF measurement tallied responses to 1,600 ($20 \times 20 \times 4$) distinct stimuli. The sheer size of the binocular stimulus set limited the stimulus configuration for each eye to a long bar that was moved along the axis perpendicular to the preferred orientation. Each stimulus was presented for three to four video frames (40–53 ms/stimulus or 19–25 stimuli/s) in a randomized sequence without any blank frames. Even at this rapid presentation rate, a complete stimulus sequence lasted 64–85 s.

Presentation of such a stimulus sequence elicited a train of spikes. For each spike generated, a causal stimulus pair that was likely to have elicited the spike was identified by looking up the stimulus that preceded the spike by a given delay (Fig. 2, top). For the real-time analysis during measurements, we used a delay of 40–70 ms, empirically determined to be effective for most neurons. However, the choice of this delay was not critical because we reanalyzed the data for delays ranging from 0 to 300 ms or more as soon as a run was completed. For example, for the rightmost spike shown in Fig. 2, a look-up in the stimulus sequence identifies k as the causal stimulus, which happens to be a pair of bright and dark bars presented to the left and right eyes, respectively, at the locations indicated. An element in the two-dimensional map at the corresponding location (X_L , X_R) is incremented. In this example, only the map for bright-dark stimulus combination is shown, but note that there are a total for four such maps, one for each permutation of bright and dark stimuli for the two eyes. Stimulus sequences were repeated, with random reshuffling of the stimulus set, until smooth profiles were obtained (or until the unit was lost). Typically, a total of 20–40 sequences was used, which took 20 min to 1 h. This process yielded a complete binocular RF map for a given correlation delay. However, as we have emphasized for the monocular case, RFs should be considered in the joint space-time domain because spatial and temporal profiles are clearly

interdependent in many cases (DeAngelis et al. 1995b). For this reason, we computed binocular RFs in the joint domain of space and time (X_L , X_R , T), or in the related domain of binocular disparity and time (D , T), where T is the correlation delay.

Histology and laminar analysis

For each electrode track, electrolytic lesions (5 μ A, 10 s) were made at 700- to 1,500- μ m intervals while the electrodes were retracted. The animal then was given an overdose of pentobarbital sodium (Nembutal) and perfused through the heart with Formalin (4% in buffered saline). Coronal sections (40 μ m thickness) of the visual cortex were made and stained with thionin. Electrode tracks then were reconstructed. Based on lesions and the depth information for each recorded cell, the laminar locations of the cells were identified. Histological analyses confirmed that all cells were recorded from area 17.

RESULTS

We recorded from a total of 257 neurons in the striate cortex of 18 normal adult cats. Of these, 115 were classified as complex on the basis of subjective criteria (Hubel and Wiesel 1962) and on the degree of temporal modulation of responses to drifting sinusoidal gratings (Skottun et al. 1991). The remaining 142 cells were classified as simple, and results for many of these cells are reported elsewhere (DeAngelis et al. 1991, 1995a; Ohzawa et al. 1996). Of the 115 complex cells, quantitative binocular RF measurements were completed for 46 disparity selective cells and 8 disparity insensitive cells. Cells were sometimes lost during preliminary grating measurements for spatial frequency tuning or dichoptic relative-phase sensitivity (Ohzawa and Freeman 1986a). We also should note that we focused on cells that showed clear relative-phase sensitivity with dichoptically presented gratings. We did not always perform complete RF mapping on cells that were nonphase-specific because they are not likely to be directly involved with stereopsis (Ohzawa and Freeman 1986b; Ohzawa et al. 1996). Therefore, our complex cell sample is biased toward those that exhibited some binocular interaction.

Figure 3A shows the measured binocular RF for a representative complex cell. There are four separate RF maps for different permutations of bright and dark stimuli presented to the two eyes. The basic characteristics of these RF maps may be summarized in three main points. First, there is a clear elongated region of strong excitation along a 45° diagonal when stimuli shown to the two eyes have the same contrast sign (i.e., bright-bright and dark-dark combinations). This pattern of response translates into a region of narrow selectivity for depth (Fig. 1C), thus eliminating many possible false matches that could occur within the cell's RF (Fig. 1D). Second, responses to bright-bright or dark-dark stimuli are nearly identical, indicating that the cell is not sensitive to the sign of contrast as long as it is matched for the two eyes. Third, for combinations of opposite stimulus contrast for the two eyes (bright-dark and dark-bright), there are two parallel regions of excitation on each side of the diagonal. These are nearly equally strong but are weaker than the excitation to the matched combinations of contrast. Note that responses near the margins of the X_L - X_R domains represent monocular excitation profiles, because one eye's stimulus is outside the RF for that eye. For example, the profile near

the left margin of each domain (i.e., the left-most vertical cross-section) represents a monocular RF profile for the right eye. Compared with the excitation level at the peak of the monocular profile, there is a suppression of the response along the diagonal for the bright-dark and dark-bright combinations. This indicates that the response is suppressed when opposite contrasts are presented at the optimal binocular disparity (for same contrast combinations). The above pattern of binocular responses is observed frequently among complex cells. These results indicate that many complex cells respond in a nearly ideal manner, filtering out responses to false targets as outlined in Fig. 1, C and D, when the sign of contrast is matched for the two eyes.

For comparison, results of a similar experiment on a simple cell are presented in Fig. 3B. The response pattern is clearly different from that of the complex cell of Fig. 3A. First, all four panels show different patterns. In particular, patterns for bright-bright and dark-dark stimulus combinations are not the same. In fact, bright-bright and dark-dark RF maps are complementary to each other: where there is a peak in one, there is a blank area in the other. This also applies to bright-dark and dark-bright conditions. The structure of the binocular RF profile for the simple cell is well predicted from the monocular response profiles given along the margins (dashed curves). The binocular pattern is approximately the sum of monocular excitation from the two eyes, i.e., those cross-points at which excitation (peaks of monocular profiles) from the left and right eyes coincide, show highly enhanced responses. On the other hand, at those locations where a trough in one eye's monocular profile meets a peak for the other eye, there is little response. This indicates that a bright-excitatory subregion is indeed inhibitory to a dark stimulus and a dark-excitatory subregion is inhibitory to a bright stimulus. This simple cell behavior is consistent with linear spatial summation of inputs from the two eyes (Ohzawa and Freeman 1986a). Note that the behavior of the complex cell (Fig. 3A) cannot be explained by linear mechanisms. The basic pattern of binocular responses shown in Fig. 3B is duplicated for other simple cells we have studied. These results suggest that simple cells are involved in general purpose processing and are not specialized specifically for stereopsis. An array of simple cells represents a general linear transformation of the retinal images. As such, they are useful for a variety of visual functions (Ohzawa et al. 1996).

Disparity tuning curves and disparity-time RFs

The format of data presentation in Fig. 3 is somewhat unusual in the sense that most binocular data in previous studies are presented in the form of a disparity tuning curve (i.e., a one-dimensional function of disparity) and not as a two-dimensional profile (Pettigrew et al. 1968; Poggio and Fischer 1977). This is primarily because of methodological limitations in previous studies. It would have taken too long with traditional peristimulus time histogram techniques, to measure responses to 400 (20 \times 20) combinations of stimulus positions. By virtue of its experimental efficiency, the reverse correlation technique largely eliminates this difficulty, allowing point-by-point measurements of binocular responses in depth. To allow comparisons, we can reduce

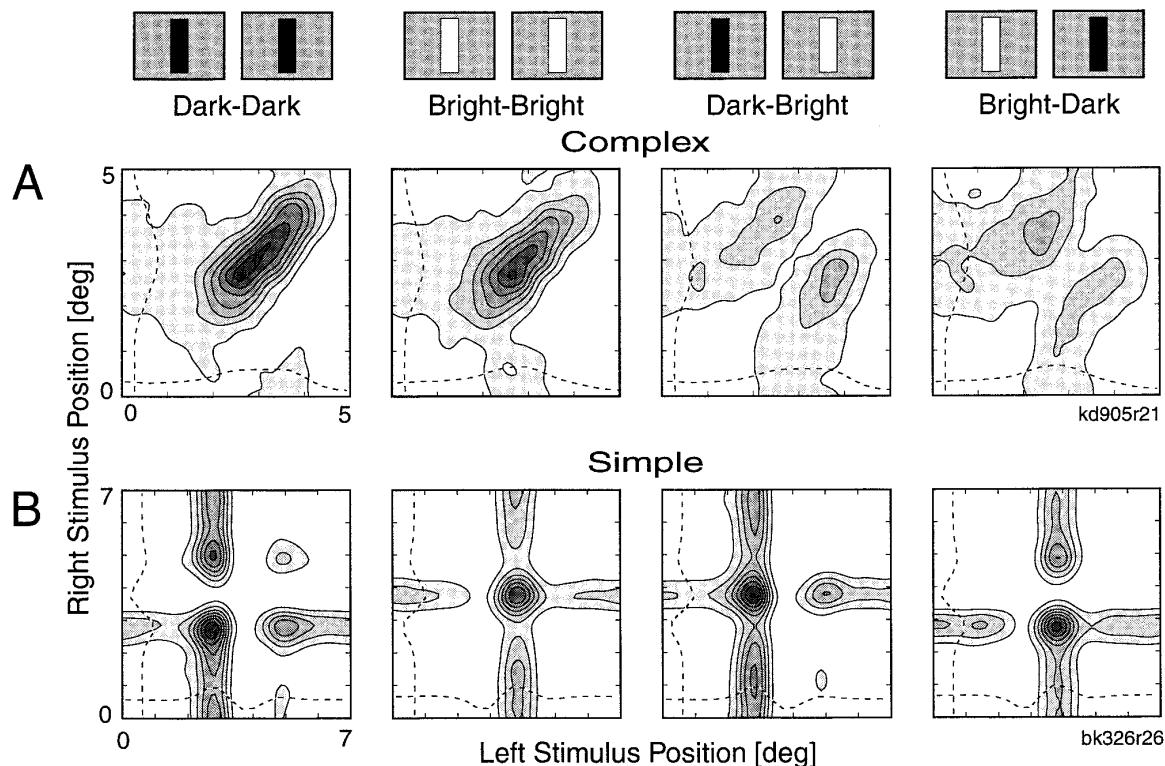


FIG. 3. Representative data from 2 cells in striate cortex are shown. *A*: complex cell (layer 2+3). Binocular RFs in X_L - X_R domain are shown as contour plots for 4 permutations of stimulus contrast. The darker the shading within a contour, the greater the response. Peak response (darkest contours) for this cell was 2.0 spikes per stimulus presentation. Contour levels are equally spaced, and the same scale is used for all 4 domains. Dashed curves near the margins of each domain represent monocular RF profiles. Stimulus size, optimal spatial frequency, and orientation were $4 \times 0.5^\circ$, 0.31 c/deg, and 75° (0° is horizontal and 90° is vertical), respectively. Stimulus duration and correlation delay were 52.8 and 65 ms, respectively. *B*: simple cell (layer 4). Monocular responses plotted near the margins of each panel (dashed curves) show the difference of bright and dark responses. Peak response for this cell was 3.9 spikes per stimulus. Stimulus size, optimal spatial frequency, and orientation were $5 \times 0.5^\circ$, 0.4 c/deg, and 155° , respectively. Stimulus duration and correlation delay were 52.8 and 60 ms, respectively.

our data to obtain disparity tuning curves. This process is illustrated in Fig. 4. A disparity tuning curve is derived by integrating the two-dimensional X_L - X_R profile along constant disparity lines parallel to the 45° diagonal. The resulting curve is shown at the top-right of Fig. 4. [Note that the pythagorean geometrical distance relationship does not hold in the X_L - X_R domain. For example, the upper left and bottom right corners of the X_L - X_R domain are separated by 10 degrees of disparity as depicted in the tuning curve, not by 7.07 degrees ($5 \cdot 2^{0.5}$). This is because the X_L - X_R domain is distorted from real space as illustrated in Fig. 1C.] Because we integrate over all combinations of positions that correspond to a particular disparity, this derivation process is essentially equivalent to measuring a disparity tuning curve with a slow moving bar stimulus, in which the total number of spikes generated by a swept bar is plotted as a function of disparity. Note that the original X_L - X_R profile is obtained using stationary flashed stimuli. Therefore, the disparity tuning curve generated in this manner is likely to be somewhat different from that obtained by swept bar stimuli. The difference will probably be more pronounced for direction selective complex cells for which second and higher order (sequential) stimulus effects make a large contribution (Baker 1990; Emerson et al. 1987). The disparity tuning curves that we obtain represent pure disparity sensitivity profiles without

the confounding effects produced by moving stimuli. Similar problems are present when attempting to interpret monocular RF profiles measured using moving bar stimuli (DeAngelis et al. 1995a; Maske et al. 1984; Ohzawa et al. 1996). Direct comparisons of our disparity tuning curves to those obtained by moving bar stimuli are not possible because we did not use such stimuli. In any case, neither our method nor the use of moving bar stimuli is appropriate for studying sequential effects. A nonlinear analysis technique (Anzai et al. 1995) is required to adequately address this problem.

In addition to facilitating comparisons between our data and those of previous studies, reducing the X_L - X_R profile to a disparity tuning curve allows us to examine the temporal behavior of disparity selectivity. We can compute disparity tuning curves (as in Fig. 4) for a range of correlation delays between stimulus and response, thus producing a disparity-time (D-T) plot. Just as the monocular space-time RF is a key predictor of direction selectivity for simple cells (Burr and Ross 1986; DeAngelis et al. 1993a,b; McLean and Palmer 1989; McLean et al. 1994; Reid et al. 1991; Watson and Ahumada 1985), the disparity-time RFs can describe how neurons respond to changes in disparity of binocular stimuli, i.e., motion-in-depth (Cynader and Regan 1978, 1982; Spileers et al. 1990). Figure 5 shows binocular RF profiles as well as disparity-time profiles for a complex cell.

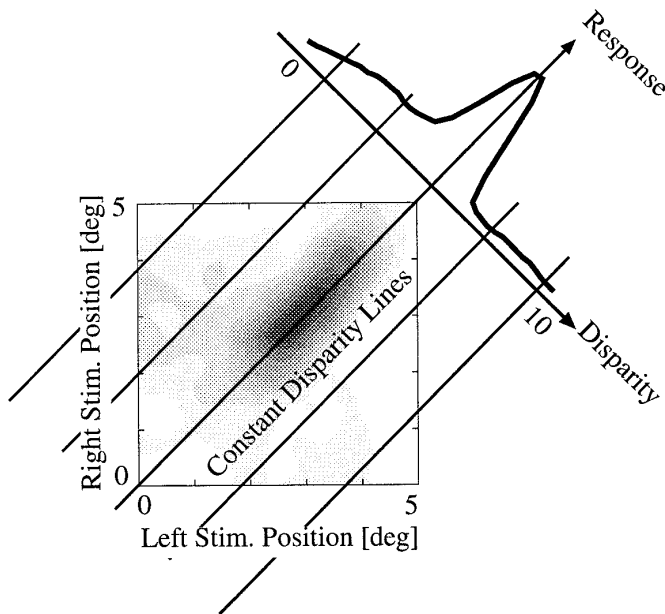


FIG. 4. A process is shown by which a traditional form of disparity tuning curve is derived from data in X_L - X_R domain. Oblique lines indicate loci along which binocular disparity of stimuli is constant. Integrating X_L - X_R data along lines of constant disparity yields the disparity tuning curve (top right).

X_L - X_R profiles (Fig. 5A) show characteristics similar to those of the complex cell in Fig. 3A. Almost identical response patterns are observed for the two matched contrast sign conditions (bright-bright and dark-dark) each having a single diagonally elongated region of excitation. The two diagonal regions of excitation for opposite contrast conditions (bright-dark and dark-bright) are also similar to those for the cell of Fig. 3A, except that the responses are stronger. Disparity tuning profiles, derived as outlined above, are shown in Fig. 5B. These profiles clearly exhibit even-symmetric disparity tuning. Figure 5C shows D-T profiles for each left-right contrast sign combination. The D-T profiles have a relatively simple structure in that the shape of the disparity tuning curve remains constant over the time course of the response. Only the response amplitude appears to change over time, with the peak response occurring at 60 ms for this cell (dashed horizontal line). Note, in particular, that the preferred disparity remains constant over time, i.e., there is no tilt of excitatory regions in the disparity-time domain. This indicates that the cell is not particularly sensitive to motion-in-depth (Cynader and Regan 1978, 1982; Spileers et al. 1990).

Figure 6 presents data from another cell. Again, X_L - X_R maps for the two matched-contrast conditions exhibit nearly identical response patterns, each having a 45° diagonal region of excitation. However, the disparity tuning curves

KD496-26

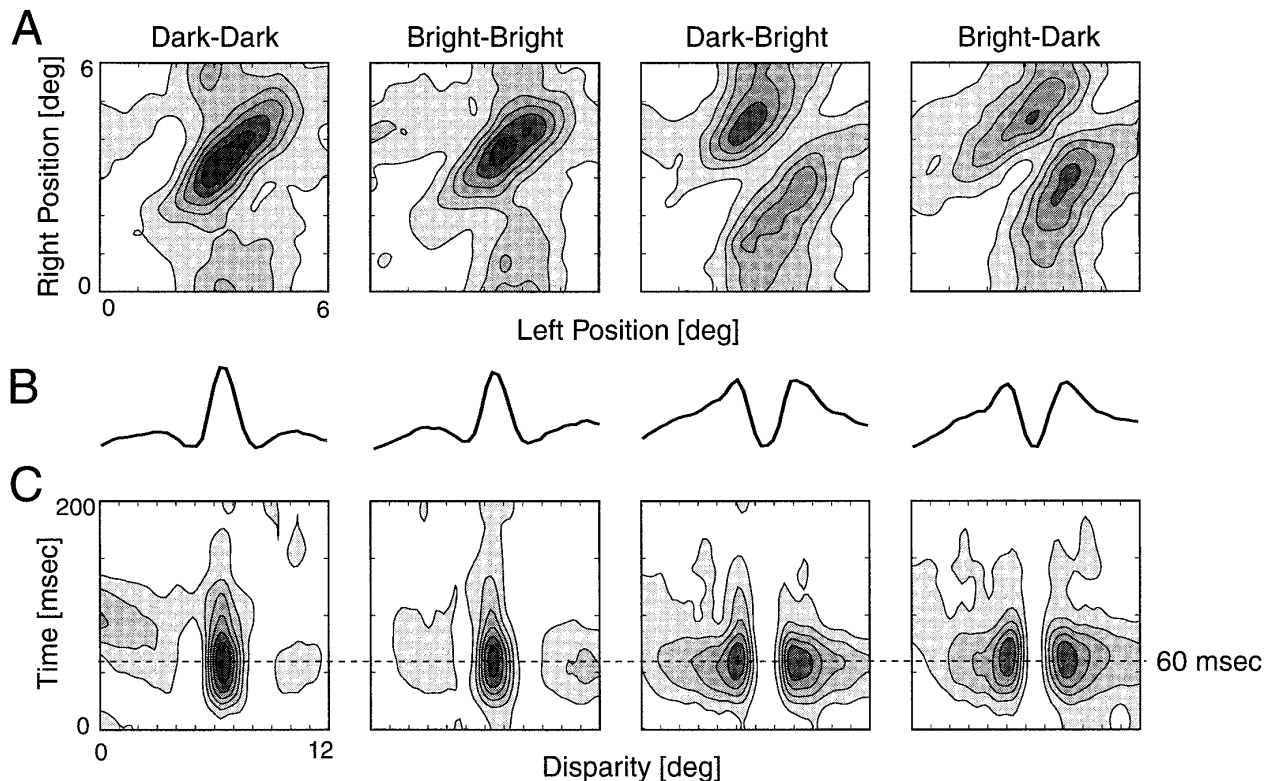


FIG. 5. Data from another complex cell (layer 2+3) are shown. A: data are presented in same format as in Fig. 3A. Peak response for this cell was 3.4 spikes per stimulus. Stimulus size, optimal spatial frequency, and orientation were $20 \times 0.4^\circ$, 0.35 c/deg, and 145° , respectively. Stimulus duration and correlation delay were 52.8 and 60 ms, respectively. B: disparity tuning curves, obtained by procedure illustrated in Fig. 4, are shown for respective X_L - X_R plots above. C: time courses of disparity tuning are shown as disparity-time (D-T) plots. Time delay used for A and B is shown by a horizontal dashed line.

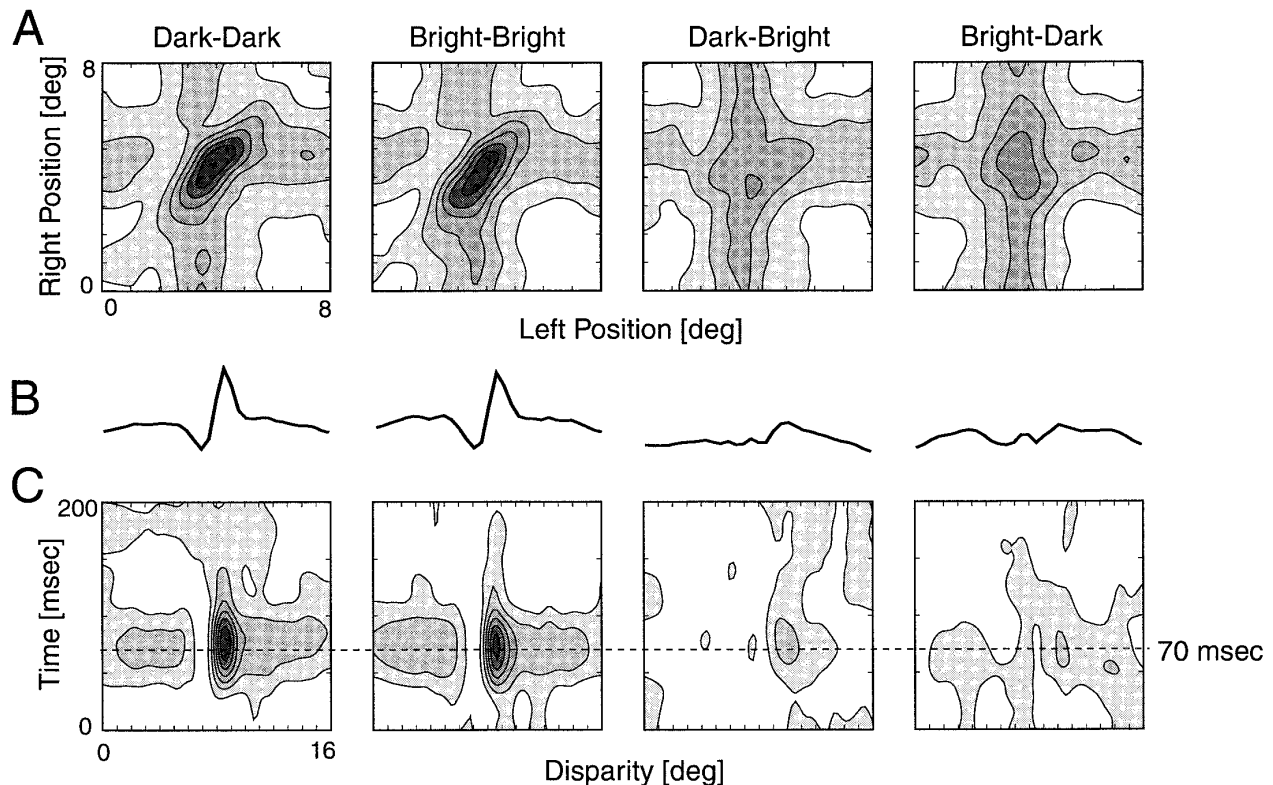


FIG. 6. Data from another complex cell (layer 2+3) are presented in same format as Fig. 5. *A*: X_L - X_R plots are shown. Peak response for this cell was 4.4 spikes per stimulus. Stimulus size, optimal spatial frequency, and orientation were $20 \times 0.4^\circ$, 0.93 c/deg, and 85° , respectively. Stimulus duration and correlation delay were 52.8 and 70 ms, respectively. *B*: disparity tuning curves for this cell are nearly odd-symmetric with respect to center of envelope. *C*: disparity-time plots are shown.

shown in Fig. 6*B* are not even symmetric. There is a prominent dip in the response below the plateau level (monocular excitation level) on the left side of the excitatory peak, but not on the right side. This disparity tuning pattern is again constant over time as indicated by the D-T profiles shown in Fig. 6*C*. Curiously, for this cell, opposite contrast conditions (dark-bright and bright-dark) produced little binocular interaction. For these conditions, the disparity tuning curves in Fig. 6*B* do not show clear peaks, and the D-T profiles in Fig. 6*C* exhibit very little structure.

Not all complex cells exhibit clear binocular interactions as shown in Figs. 3, 5, and 6. Figure 7 presents results from a cell that had no apparent binocular interaction. In Fig. 7*A*, excitation due to left and right eye monocular stimuli extends as vertical and horizontal bands, respectively, forming a cross-shaped profile. At the intersection of these bands, the excitation level is generally higher. This is most pronounced for the left-most panel (dark-dark), but it is also visible for the middle two panels. This pattern occurs because excitation from the two eyes adds together at the intersection. However, the transformed disparity tuning curves are flat, as shown in Fig. 7*B*. This may be understood by examining the procedure for deriving a disparity tuning curve, as depicted in Fig. 4. A diagonal line at any location in Fig. 4 will cross the peaks of excitation for the left and right eyes, given a sufficient path length. Therefore the cumulative response remains constant regardless of whether the monocular exci-

tation coincides or occurs at separate positions along a given path. The existence of these cells is expected, because complex cells that are not selective to disparity have been reported previously (Chino et al. 1994; Joshua and Bishop 1970; Ohzawa and Freeman 1986b). The cell shown in Fig. 7 did not exhibit interocular phase tuning to drifting sinusoidal grating stimuli presented dichoptically (Ohzawa and Freeman 1986b).

Disparity energy model

We have proposed a model for disparity-selective complex cells based on a combination of simple-cell subunits (Ohzawa et al. 1990). Here, we present the model's behavior in detail and examine quantitatively whether it provides an adequate description of experimental data. Figure 8*A* illustrates the model for a complex cell that is tuned to zero disparity. The model consists of a minimum of four simple-cell subunits that are combined to produce the output of a complex cell. In this sense, the model is hierarchical and gives a concrete functional design to the original scheme proposed by Hubel and Wiesel (1962). Our model is also a natural binocular extension of monocular complex cell models (Adelson and Bergen 1985; Emerson et al. 1992; Pollen et al. 1989). Each subunit is binocular and linearly combines inputs from the two eyes (Ohzawa and Freeman 1986a). The output of the subunits passes through a half-squaring nonlinearity before

KD458-34

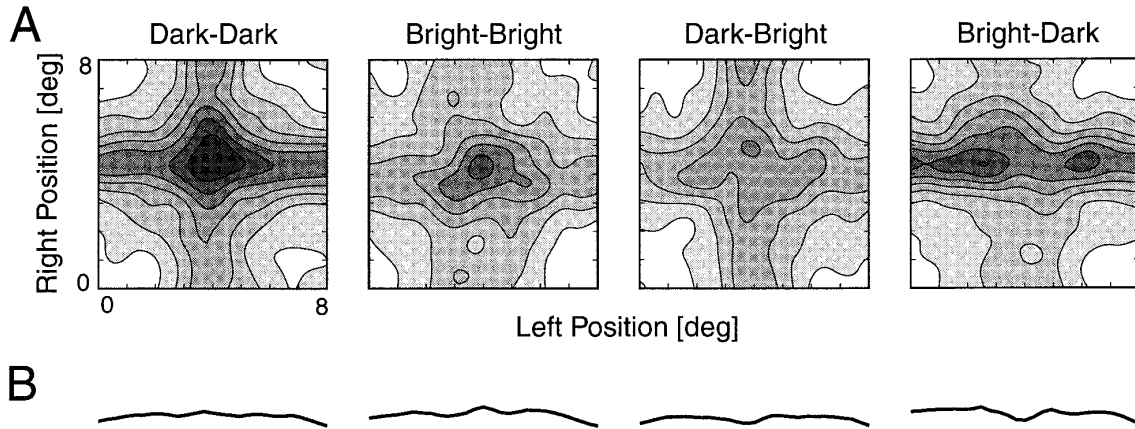


FIG. 7. Data are illustrated for a complex cell (layer 5) that is not sensitive to binocular disparity. *A*: there is no diagonal structure in these X_L - X_R plots unlike cells in previous figures. Peak response for this cell was 2.7 spikes per stimulus. Stimulus size, optimal spatial frequency, and orientation were $20 \times 0.25^\circ$, 0.95 c/deg, and 160° , respectively. Stimulus duration and correlation delay were 52.8 and 100 ms, respectively. *B*: disparity tuning curves are nearly flat for all permutations of stimulus contrast. D-T plots are not shown because there is no structure in any of the plots.

converging onto a complex cell (Emerson et al. 1989, 1992; Heeger 1992a,b; Ohzawa et al. 1990; Pollen et al. 1989). The output of the subunit S1 at the top of Fig. 8A to a line stimulus (or a thin bar) is given by

$$R_{S1}(X_L, X_R) = \text{Pos}[\exp(-kX_L^2) \cos(2\pi fX_L) + \exp(-kX_R^2) \cos(2\pi fX_R + \psi)]^2 \quad (1)$$

where k is the factor that determines the width of the subunit RFs and f is the spatial frequency. These two parameters are assumed to be equal for the two eyes, and this assumption is supported by our data for simple cells (Ohzawa et al. 1996). The parameter ψ is the phase difference between left and right RFs. $\text{Pos}[v]$ is a half-rectifying function

$$\begin{aligned} \text{Pos}[v] &= v \quad \text{for } v \geq 0 \\ &= 0 \quad \text{for } v < 0. \end{aligned} \quad (2)$$

This is a relatively straightforward model of a simple cell with linear binocular convergence (Ohzawa and Freeman 1986a) and a half-squaring nonlinearity (Emerson et al. 1989, 1992; Heeger 1992a,b; Pollen et al. 1989). This latter component may be considered as a form of “soft threshold” (Carandini et al. 1996). The phase difference, ψ , accounts for the observation that the left and right RFs of simple cells may have different shapes (DeAngelis et al. 1991, 1995a; Freeman and Ohzawa 1990; Ohzawa et al. 1996). For simplicity, we use even- and odd-symmetric subunit RFs as shown in Fig. 8A. This nonbiological restriction (DeAngelis et al. 1993a; Field and Tolhurst 1986) is removed later.

In Fig. 8A, we show the case of $\psi = 0$. Plots for $\psi \neq 0$ are shown later. Because the subunits S1 and S2 have inverted RFs and for any function $g(a)$

$$\text{Pos}[g(a)]^2 + \text{Pos}[-g(a)]^2 = g(a)^2, \quad (3)$$

the sum of contributions from the top two subunits S1 and S2 is given by

$$R_{S1}(X_L, X_R) + R_{S2}(X_L, X_R) = \{\exp(-kX_L^2) \cos(2\pi fX_L) + \exp(-kX_R^2) \cos(2\pi fX_R + \psi)\}^2 \quad (4)$$

The response of complex cell C1 is then given by adding the contribution from subunits S3 and S4 to Eq. 4

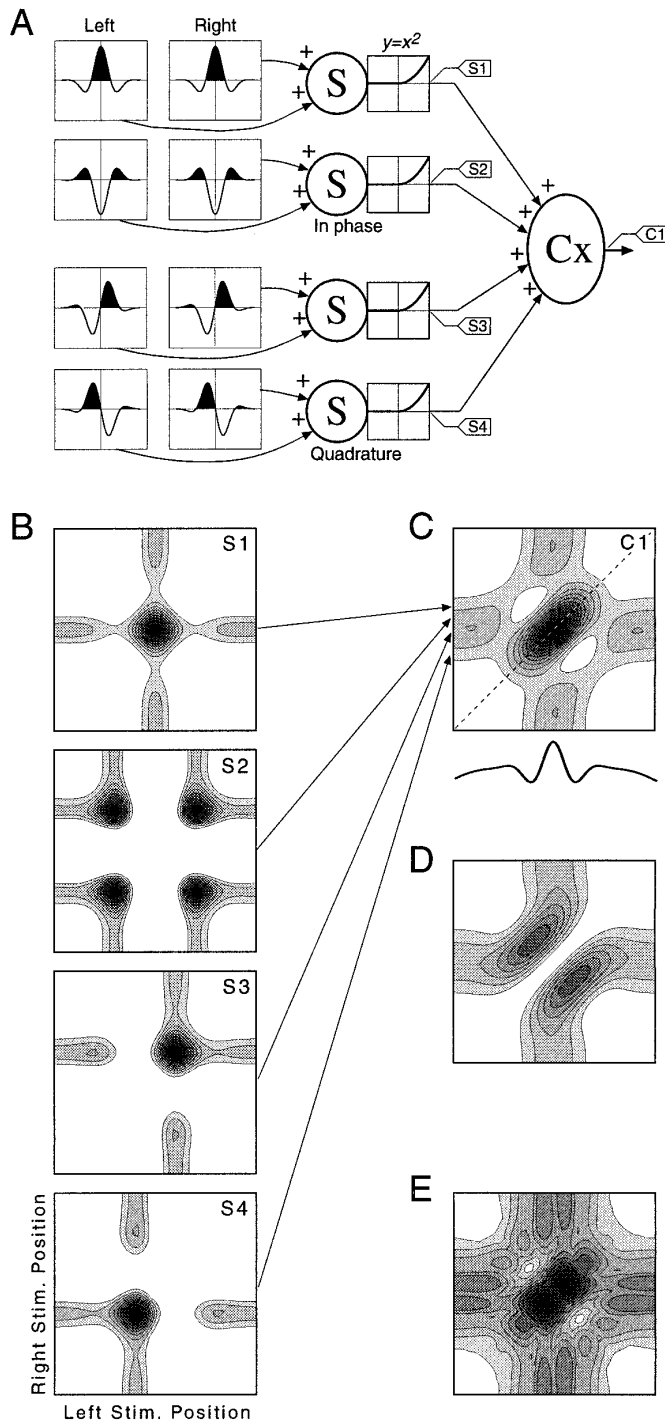
$$\begin{aligned} R_{C1}(X_L, X_R) &= R_{S1} + R_{S2} + R_{S3} + R_{S4} = \{\exp(-kX_L^2) \cos(2\pi fX_L) \\ &+ \exp(-kX_R^2) \cos(2\pi fX_R + \psi)\}^2 + \{\exp(-kX_L^2) \sin(2\pi fX_L) \\ &+ \exp(-kX_R^2) \sin(2\pi fX_R + \psi)\}^2. \end{aligned} \quad (5)$$

This function is shown as a contour plot in Fig. 8C and is very similar to the responses of the complex cells shown in Figs. 3A and 5A for the matched contrast conditions. A traditional disparity tuning curve also is obtained for the model, using the procedure described in Fig. 4, and plotted below the two-dimensional profile in Fig. 8C. The disparity tuning curve is quite similar to those shown in Figs. 4 and 5B. The case described in Fig. 8A ($\psi = 0$) produces a cell that is tuned to zero disparity as indicated by the dashed diagonal line in Fig. 8C.

Models that compute the sum-of-squares (e.g., Eq. 5) are called “energy models” on the basis of a formal definition of energy (Adelson and Bergen 1985). For example, in physics, the integral over time of the square of a voltage waveform across a resistor is proportional to the energy dissipated within the resistor. This notion may be generalized to neural signals. Simple-cell subunits that feed into a complex cell (i.e., a binocular energy unit) must meet specific requirements to produce a sufficiently smooth binocular profile. First, all monocular parameters must be the same among the four subunits, including spatial frequency, orientation, size, and position of the RF envelopes. This has been shown to be true for most simple cells (Ohzawa et al. 1996; Skottun and Freeman 1984). The requirement does not apply to phase, however. Second, all subunits must share a common preferred disparity as measured with bar or grating stimuli (Ohzawa and Freeman 1986b). Although simple cells do not possess a unique preferred disparity to noise stimuli and hence they tend not to respond to dynamic noise stereograms (Poggio et al. 1985; Qian 1994; Zhu and Qian 1996), they do exhibit a clear disparity tuning to bar or grating stimuli

(Ferster 1981; LeVay and Voigt 1988; Ohzawa and Freeman 1986b). Third, the phases of the four subunits (not left and right RF phases) must differ from each other by multiples of 90° (Pollen et al. 1989). Our model assumes that these conditions are fulfilled.

To illustrate how each of the four subunits contributes to the final complex cell response, X_L - X_R maps for the individual subunits are shown in Fig. 8B. The binocular response profile of each subunit exhibit the pattern of excitation and inhibition that is well predicted by monocular RFs of the subunit. For example, even-symmetric monocular RFs of the



subunit at the *top* of Fig. 8A yield the binocular response at S1, which is even-symmetric along both the horizontal and vertical axes. The response is strongest at the center of the profile, where the peaks of monocular excitation coincide. Responses of the second subunit, at S2, are shown in the second panel of Fig. 8B. Again, the binocular response pattern is related closely to the monocular RF structure shown in Fig. 8A (second subunit from top). The binocular response exhibits four peaks at locations where peaks in the two monocular RF profiles coincide. Note the similarity of the binocular responses of the model's subunits to those of the simple cell shown in Fig. 3B. This is not surprising given that binocular simple cells combine input from the two eyes in a linear manner (LeVay and Voigt 1988; Ohzawa and Freeman 1986a), as do the subunits in the model. The binocular response of the complex cell, as an X_L - X_R map, is given in Fig. 8C. This is the sum of the four subunit profiles shown in Fig. 8B. Note that the four subunit profiles, each of which lacks any elongation along the diagonal, combine to produce a remarkably smooth, diagonally elongated complex cell binocular RF profile.

Binocular responses to opposite contrast stimuli to the two eyes are shown in Fig. 8D (Ohzawa et al. 1990). The response pattern may be expressed as

$$R_{C1}(X_L, X_R) = \{ \exp(-kX_L^2) \cos(2\pi fX_L) - \exp(-kX_R^2) \cos(2\pi fX_R + \psi) \}^2 + \{ \exp(-kX_L^2) \sin(2\pi fX_L) - \exp(-kX_R^2) \sin(2\pi fX_R + \psi) \}^2. \quad (6)$$

This equation is identical to Eq. 5, except for the sign inversion of the right-eye terms. There is prominent suppression along the diagonal, exactly at the location where there is an excitatory region in Fig. 8C. In this case, the cell exhibits two diagonal bands of excitation for opposite contrast conditions, again at the disparities where there was suppression in Fig. 8C.

The form of nonlinearity that follows binocular convergence in the linear subunits is important. The model shown in Fig. 8A employs a squaring nonlinearity that produces a smooth binocular response profile as shown in Fig. 8C. If, instead, a simple half-wave-rectifying nonlinearity is used, the otherwise identical model produces a binocular response as shown in Fig. 8E. For both monocular and binocular portions of the responses (responses near the edges and the

FIG. 8. A disparity energy model is illustrated. A: a complex cell (Cx) tuned to 0 disparity is modeled as consisting of 4 simple subunits (S). Each subunit combines input from 2 eyes linearly according to left and right RFs (*left*). Output of each subunit goes through a half-squaring nonlinearity that represents the fact that only postsynaptic potentials that exceed a threshold value elicit action potentials. "Tap points" (S1-S4) are included for reference in B. B: binocular responses in X_L - X_R domain are shown as contour plots for 4 subunits. The darker the shading, the stronger the response. Responses are normalized for each plot to show details of weaker responses such as those for S2. C: binocular responses of complex cell are shown. This is simply a point-by-point sum with appropriate scaling of all 4 subunit RFs in B. The plot represents responses to bright or dark bar stimuli presented to both eyes. Compare this to experimental data in Fig. 3A. D: binocular responses to opposite contrast conditions are shown, i.e., a bright bar to 1 eye and a dark bar to the other. E: responses are shown of a model that employs a half-rectifier instead of a half-squarer as the output nonlinearity for simple subunits. Note that substantial ripples remain in the complex cell response profile, but the overall response pattern is similar to that of C.

central diagonal, respectively), substantial ripples are observed in the profile. Except for these ripples, however, the basic pattern of the response is similar to that of the squaring configuration shown in Fig. 8C. Additional subunits may be used in the model to smooth out the ripples to obtain a final smooth profile. However, the nonlinearity must be of a squaring form if the number of subunits is to be minimized.

Phase model

We previously have proposed an efficient scheme for encoding binocular disparity information by a population of simple cells in the striate cortex (DeAngelis et al. 1991, 1995a; Freeman and Ohzawa 1990; Ohzawa et al. 1996). This scheme is called a phase model because disparity information is encoded by binocular simple cells that have different RF phases (i.e., shapes) for the two eyes. The traditional notion of binocular disparity encoding is based on a position model, in which disparity is encoded via positional offsets of left and right RFs. We now examine predictions of our complex-cell model whose subunits have phase or position shifts. If the model provides different results for the phase and position schemes, we will be able to evaluate the schemes by examining data from complex cells.

A predicted response of a binocular energy unit based on the phase model is shown in Fig. 9A. Phases of the subunit RFs (*right column*) differs by 90° for the two eyes. We previously have shown that a substantial fraction of simple cells have RF profiles that are different for the two eyes (DeAngelis et al. 1991, 1995a; Freeman and Ohzawa 1990; Ohzawa et al. 1996). Therefore the scheme shown in Fig. 9A is quite reasonable with respect to the existence of the required subunits. Note that the only difference between the model complex cells in Fig. 8A and 9A is that the phases of the right eye RFs are shifted by 90° in the same direction for all four subunits. With this construction, the binocular response (Fig. 9A, *left*) no longer shows a symmetric profile with respect to the diagonal (compare with Fig. 8C). Instead, a region of excitation is shifted from and parallel to the 45° diagonal indicated by a dashed line. The odd-symmetry of the binocular response is clear in the disparity tuning curve shown below the two-dimensional profile in Fig. 9A. This response pattern is very similar to that of the cell shown in Figs. 6, A and B (*left 2 panels*). It will be shown below that the phase of this disparity tuning curve is ψ , the phase difference between the left and right RFs of the subunits.

Note that the intersection of the peaks of monocular excitation (intersection of vertical and horizontal dashed lines) does not coincide with the peak of the binocular response, indicated by the contours with the darkest shading. In other words, a combination of monocularly optimal stimuli does not result in a binocularly optimal stimulus. The maximum binocular response is obtained when stimuli for the two eyes are at nonoptimal locations within the monocularly measured RFs. This may be how a binocular complex cell that prefers a non-0 disparity is constructed from subunits that have different RF structures for the two eyes, and it provides an extension of the phase model to complex cells (Fleet et al. 1996, 1997).

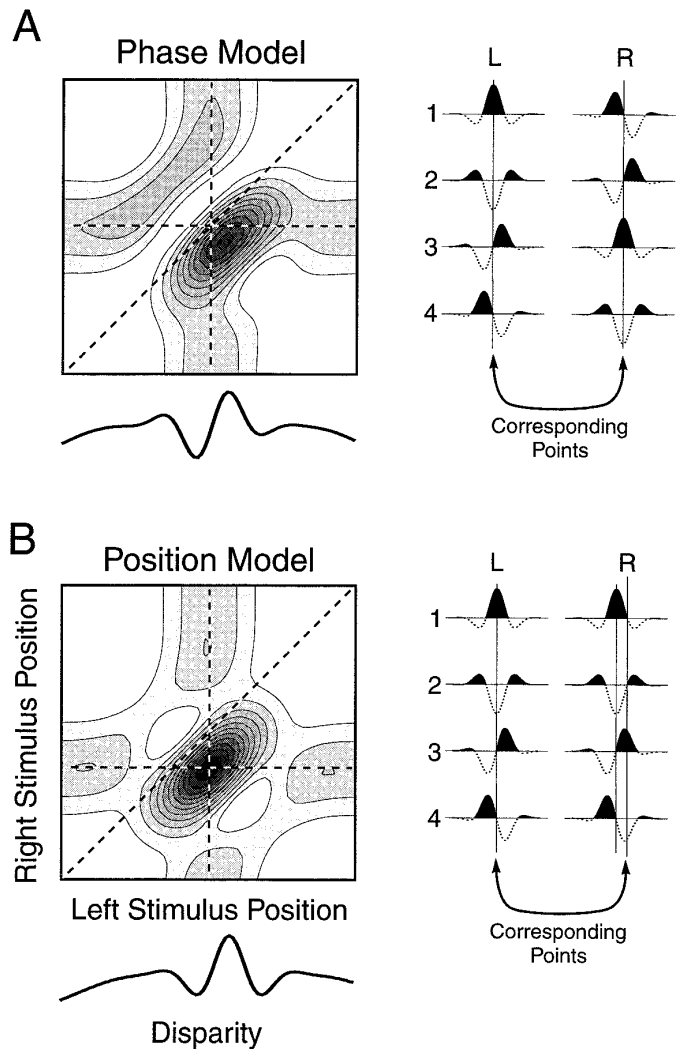


FIG. 9. Two possible mechanisms are shown for creating complex cells that are tuned to a non-0 binocular disparity. A: phase model. Four subunits (*right*) have RF structures that differ in phase. All 4 units have a left-right phase difference of 90° . *Left*: vertical and horizontal dashed lines represent peaks of monocular RF profiles for left and right eyes, respectively. Peak binocular response (center of darkest contour) does not fall on diagonal dashed line. Disparity tuning curve (solid curve below X_L - X_R plot) is odd-symmetric. B: position model. Organization of subunits is identical to that of Fig. 8A, except that all subunits have a positional shift of their right RFs. The X_L - X_R plot is also a shifted version of Fig. 8C. Disparity tuning curve is even-symmetric.

Position model

The position model combines four simple-cell subunits exactly as in Fig. 8A, except with a common positional offset. This is shown in Fig. 9B (*right*). Because the underlying structure of the subunits is the same as that for Fig. 8A, the binocular response pattern is also the same. The whole response pattern is shifted downward, reflecting the offset of the right RFs of the subunits. The disparity tuning curve obtained from the two-dimensional map has exactly the same shape as that for Fig. 8C and is even symmetric. The cell has a preferred disparity that is non-0 as indicated by the shift of the excitatory peak from the diagonal dashed line. The position of the peak of the binocular response is completely predictable from the intersection of the peaks of

monocular responses, as shown by the fact that the peak of the binocular response lies exactly at the crossing of the horizontal and vertical dashed lines. Note that the binocular energy units shown in Figs. 9B and 8A have identical shapes, despite the fact that the units are tuned to different disparities. Thus these two cases cannot be distinguished unless one has an accurate measurement of zero disparity (i.e., corresponding points). This information is not readily available in our paralyzed preparation.

Hybrid phase and position model

It is possible that neither the phase nor position model accurately describes the binocular RFs of cortical cells. Disparity information may be encoded by a hybrid mechanism based on both phase and positional differences (Anzai et al. 1995; Fleet et al. 1996, 1997; Jacobson et al. 1993; Qian and Zhu 1997; Zhu and Qian 1996). Preliminary evidence from our lab indicates that, in fact, both phase and positional offsets contribute to a simple cell's disparity preference (Anzai et al. 1995). In the absence of absolute eye position information in our paralyzed preparation, the positional offset component cannot be measured. However, the phase component may be estimated exactly by determining the phase of the disparity tuning curve. In other words, for the energy model, the phase of the disparity tuning profile exactly reflects the phase difference, ψ , of the RF profiles for the two eyes, regardless of the degree of positional offset. Therefore, the symmetry of binocular response profiles provides a signature for the phase model. However, the hybrid model cannot be ruled out even if asymmetry is found in the profiles.

Evaluation of the models

With the limitations described above, we are able to evaluate the validity of the models. Specifically, we can determine experimentally the contribution of the phase difference between left and right subunit RFs to a given complex cell's disparity tuning. This is because there is a direct relationship between the left-right phase difference of the subunit RFs and the symmetry of the binocular response profile for the complex cell that combines these subunits. We have seen this graphically for the cases of phase difference $\psi = 0^\circ$ (Fig. 8) and $\psi = 90^\circ$ (Fig. 9A). In general, for any value of ψ , the expression in Eq. 5 may be simplified as

$$R_{CI}(X_L, X_R) = \exp(-2kX_L^2) + \exp(-2kX_R^2) + 2 \exp[-k(X_L^2 + X_R^2)] \cos[2\pi f(X_L - X_R) - \psi] \quad (7)$$

Equation 7 shows that the binocular response of a disparity energy unit defined by Eq. 5 may be expressed as the sum of three terms: two monocular response terms and a binocular term (Fleet et al. 1996, 1997). The monocular terms describe Gaussian-shaped profiles as would be obtained by monocular mappings of the left and right RFs. The third (binocular) term is a two-dimensional Gabor function (Daugmann 1985; Gabor 1946; Jones and Palmer 1987a,b; Marceljā 1980) that is oriented at 45° . This relationship is

described graphically in Fig. 10A for the case of $\psi = 0^\circ$. It is clear from Fig. 10A that the diagonal region of excitation in the *leftmost panel* is due to the binocular term plotted in the *rightmost panel*. Monocular responses (*middle 2 panels*) appear horizontal and vertical bands of excitation and determine the response near the margins of the *leftmost panel*. Two suppressive bands (shown with dashed contours in the *rightmost panel*) on each side of the excitatory diagonal region create interruptions in the binocular response in the *leftmost panel*. For an intuitive interpretation of this binocular response (the *rightmost panel*) in actual three-dimensional space, it may be helpful to visualize a sandwich held out in space, with layers of the sandwich perpendicular to the cyclopean direction of gaze. A stimulus of appropriate orientation will be excitatory if it falls within the central (filling) layer and inhibitory if it falls within the surrounding (bread) layers.

Note that the monocular terms are independent of the phase difference, ψ , between left and right RFs, and that the symmetry of the binocular response profile is determined solely by the third (binocular) term of Eq. 7. In Fig. 10B, this binocular term is shown for the phase difference $\psi = 90^\circ$. One may see readily how the pattern of response presented in the *leftmost panel* is produced by the sum of the monocular responses and the binocular term (the *rightmost panel*), which, here, is odd symmetric.

Observe also that the subunit RFs do not have to be even or odd symmetric as shown in Fig. 8A, i.e., their phases do not have to be multiples of 90° . The absolute phases of subunit RFs do not affect the complex cell response because they are canceled in the third term of Eq. 7 (Fleet et al. 1996; Qian 1994; Zhu and Qian 1996). Only the phase difference between subunits (not ψ , the phase difference between left and right RFs) must be multiples of 90° for a minimum configuration. If more subunits are allowed, even this quadrature constraint may be removed.

To examine how well the energy model fits the data from cells, we have performed decompositions of binocular responses from the cells shown in Figs. 5 and 6. Results are illustrated in Fig. 11. Figure 11A shows a decomposition of the binocular response (dark-dark) for the cell of Fig. 5. The original data in the *leftmost panel* may be decomposed into left eye, right eye, and binocular responses in the three panels to the *right*. The *rightmost panel* shows the residual error of the fit. Fitting of the two-dimensional profile was performed by a modified Levenberg-Marquardt optimization algorithm using Matlab (MathWorks). To allow for variations in the data such as inexact centering of RFs, ocular dominance (Hubel and Wiesel 1962), and other scaling factors, the actual function used for the fit is given by

$$A_L \exp[-2k(X_L - X_{OL})^2] + A_R \exp[-2k(X_R - X_{OR})^2] + 2A_B \exp[-k((X_L - X_{OL})^2 + (X_R - X_{OR})^2)] \times \cos[2\pi f((X_L - X_{OL}) - (X_R - X_{OR})) - \psi] + C \quad (8)$$

where the additional parameters are as follows: X_{OL} and X_{OR} are center positions of monocular RFs for the left and right eyes. A_L , A_R , and A_B are scaling factors that account for ocular dominance and balance of monocular and binocular terms. Note that, for a strict energy model, $A_B = A_L \cdot A_R$. The spatial frequency of the two-dimensional Gabor function

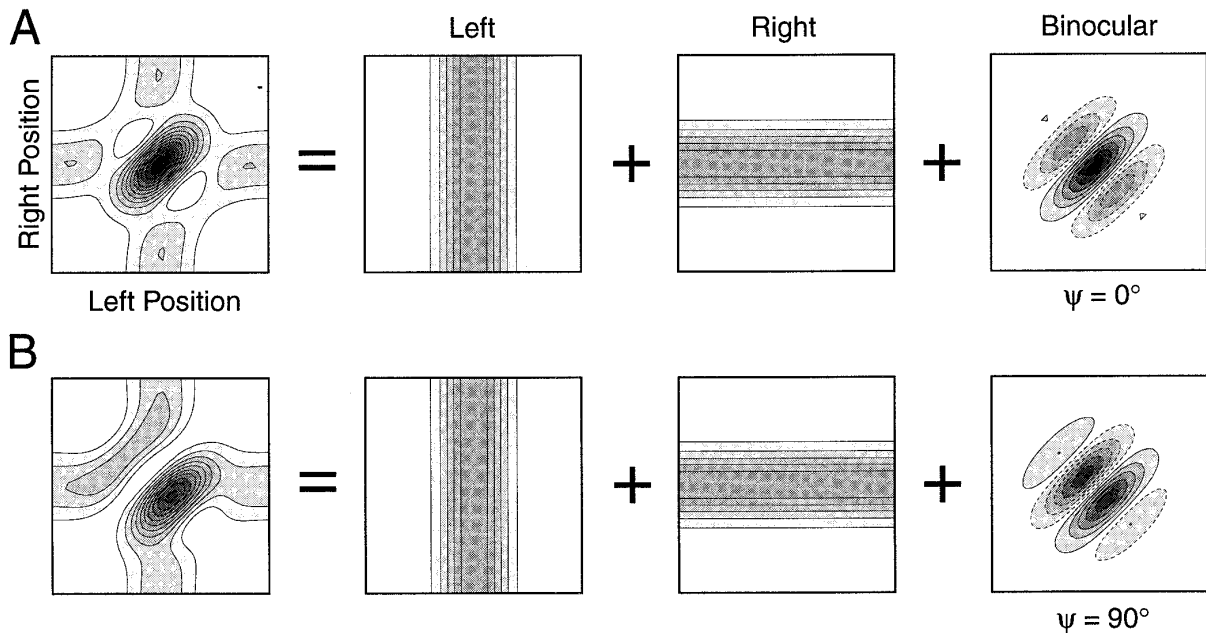


FIG. 10. Responses of complex cell energy model in X_L - X_R domain may be decomposed into 3 terms: left and right monocular response terms and a purely binocular term. This decomposition is shown graphically for 2 phase differences of left and right subunit RFs, A: $\psi = 0^\circ$ and B: $\psi = 90^\circ$. Monocular components are identical and independent of phase difference. Binocular term is a Gabor function that is oriented at 45° . Phase difference, ψ , determines symmetry of binocular component.

for the binocular term is given by f , and C is a constant offset that may be necessary to account for spontaneous discharge and spikes that are uncorrelated to stimuli. A total of nine free parameters is used. Three panels that represent the monocular and binocular terms of the fit, and a fourth that shows the residual error, are plotted (with the same scale) in Fig. 11A. The fit appears reasonable as there is no highly systematic structure in the error profile. The right eye response is substantially weaker than the left eye response. The binocular response component shows nearly exact even-symmetry ($\psi = 3.7^\circ$, see Fig. 10) with an excitatory region at the center (Fig. 11A). The intersection of the vertical and horizontal dashed lines, representing the peak positions of left and right monocular excitation, respectively, falls exactly on the peak of the binocular component. Recall that this was also the predicted behavior of the model for $\psi = 0^\circ$, as shown in Fig. 8. A fit for the bright-bright data gives nearly an identical set of parameters for the binocular component ($\psi = -7.0^\circ$, not shown). Figure 11B shows a decomposition of the bright-dark response from the same cell. Monocular responses are quite similar to those in Fig. 11A, indicating insensitivity to the sign of stimulus contrast. However, the binocular response shows a clear inversion of phase as indicated by a suppressive central region shown by the dashed contours ($\psi = 187.4^\circ$; see Fig. 8D for comparison with the model). Again, there is no obvious structure in the error profile, indicating a reasonably good fit. Overall, the data for this cell are represented well by the disparity energy model of Eqs. 5 and 7.

Data from another cell (shown previously in Fig. 6) are fit by the same procedure, and the results are shown in Fig. 11, C and D. This cell had an asymmetric disparity tuning curve (Fig. 6B). This also is revealed in the binocular re-

sponse term (the 4th panel in Fig. 11C) that shows a nearly odd-symmetric profile ($\psi = 70^\circ$). The parameters of the fit to the bright-bright responses are also similar ($\psi = 82^\circ$, not shown). In agreement with the predictions shown in Fig. 9B, the intersection of the peaks of the monocular RFs (dashed lines in the 1st and 4th panels in Fig. 11C) does not coincide with the peak of the fit to the binocular response (Ohzawa et al. 1990). Thus the energy model, as shown in Fig. 9A, provides a reasonable fit. However, there are significant deviations from predictions of the energy models. The residual error of the fit in Fig. 11C shows clear structure in the form of regions that are oriented approximately at 45° . It appears as if the positive diagonal region in the Gabor function for the binocular response does not provide a sufficiently long diagonal to fit the data. In addition, responses to opposite contrast conditions for the two eyes, shown in Fig. 11D (and Fig. 6), exhibit hardly any diagonal structure. This is reflected in the small amplitude of the binocular response term (the 4th panel of Fig. 11D). Clearly, this is a deviation from the predictions of the simplest energy model presented in Eqs. 5 and 7.

Monocular analyses of complex cell RFs provide a clue to the possible cause of the residual error shown in Fig. 11C. It has been shown that a single energy unit (consisting of 4 subunits as shown in Figs. 8 and 9) may not necessarily cover the whole RF area of the complex cell, and therefore multiple (at least ≤ 4 or 5) energy units are needed for some complex cells (Ohzawa et al. 1995). Details of this analysis will be presented elsewhere. Note that the spatial extent of an energy unit cannot be increased by simply using a larger subunit because this will increase the spatial extent of the binocular term uniformly in all directions, not simply along the 45° diagonal (see DISCUSSION). In principle, it should be

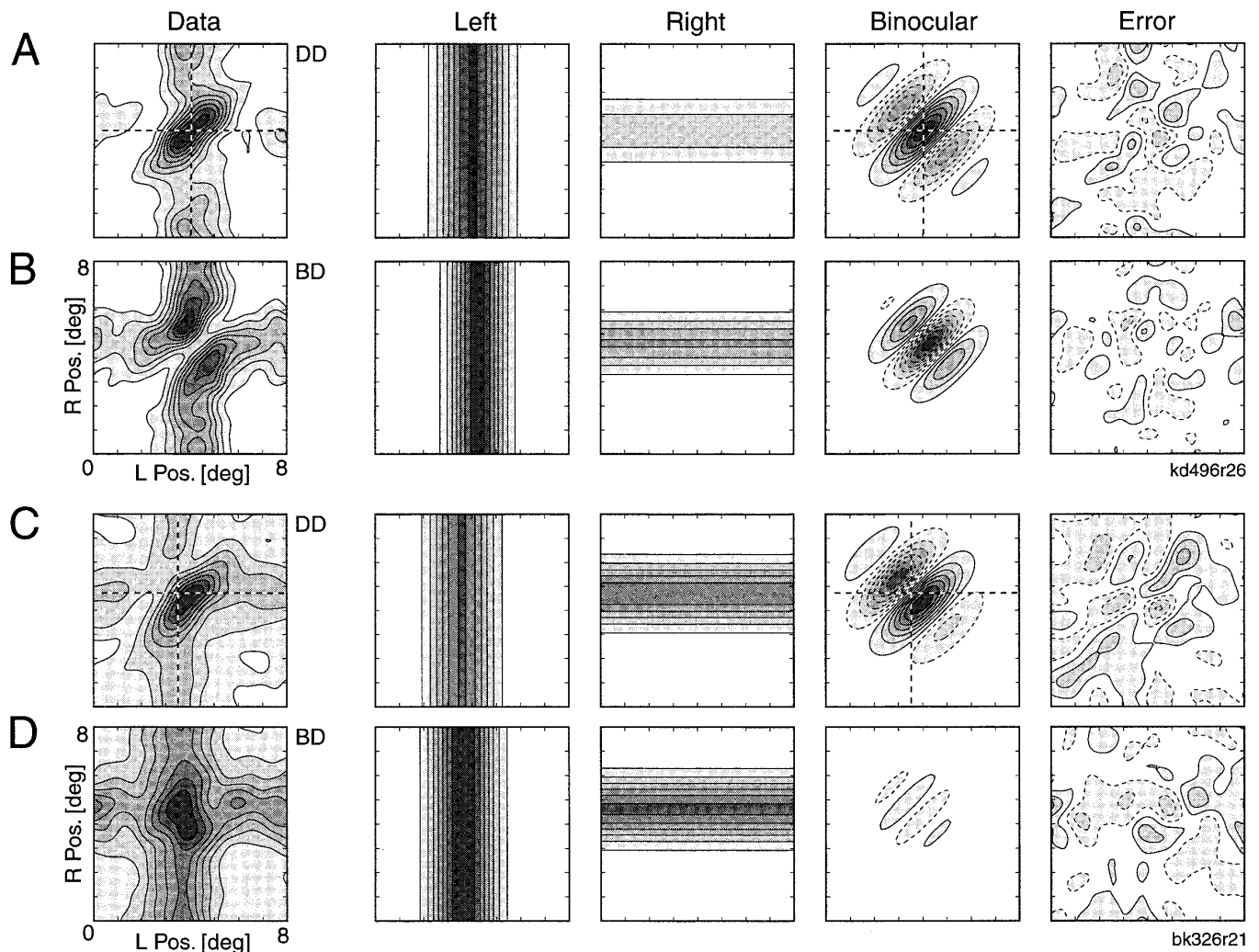


FIG. 11. Data from 2 complex cells are fit by the energy model. Two monocular terms and a binocular term of fit (see Eq. 8) are shown along with the residual error term. All 4 terms of fit are presented as contour plots using the same amplitude scale. A: fitted profile (dark stimuli for both eyes) for complex cell presented in Fig. 5 is shown. Data are from an additional measurement performed using a larger spatial domain ($8 \times 8^\circ$) than that shown in Fig. 5 ($6 \times 6^\circ$). Vertical and horizontal dashed lines in data and binocular term panels indicate peak locations of left and right monocular responses, respectively. B: same as A except that responses to bright-dark stimuli are shown. C: fitted profile (dark stimuli for both eyes) is shown for the cell presented in Fig. 6. D: same as C except that responses to bright-dark stimuli are shown.

possible to model these profiles by allowing multiple energy units (each with 4 simple-cell subunits) and fitting the model to the data. However, the use of multiple energy units presents a practical problem in data modeling, because it is likely to cause a loss of stability and uniqueness of solutions due to the increased number of free parameters. Therefore fits based on multiple energy units have not been attempted. With these limitations in mind, we have proceeded with fits of the single energy unit model to data.

Fits of a single energy unit are shown in Fig. 12 for four more cells from this group. Only fits for dark-dark and dark-bright responses are shown for each cell, because profiles for the remaining two conditions of contrasts are generally very similar. In addition, the fits are not broken down to three terms (see Eq. 7), but instead are shown as the sum of all terms. For all four examples, the fits capture the main features of the raw data very well for both dark-dark and dark-bright conditions. The residual error of the fits is small

(fit and error panels have the same amplitude scale in each case), and no systematic structure is apparent. Note that Fig. 12D depicts a cell with a large imbalance in response strengths for bright and dark stimuli, with the bright stimulus eliciting only weak excitation. This is reflected in the fact that the data and fit panels for the dark-bright condition (Fig. 12D, right) consist predominantly of vertical contours that show a dominant left eye response. Bright response for the right eye is very weak as indicated by the absence of horizontal contours. The results from these cells confirm that a single energy unit model is sufficient to describe binocular behavior of a substantial subset of complex cells.

However, this model is clearly not adequate for some cells and two examples of this case are shown in Fig. 13. In Fig. 13A, a cell with a highly elongated diagonal is illustrated, and this cannot be fit well with a single energy unit model. The error term for the dark-dark condition (Fig. 13A, left) shows clear high amplitude residual peaks that lie on the

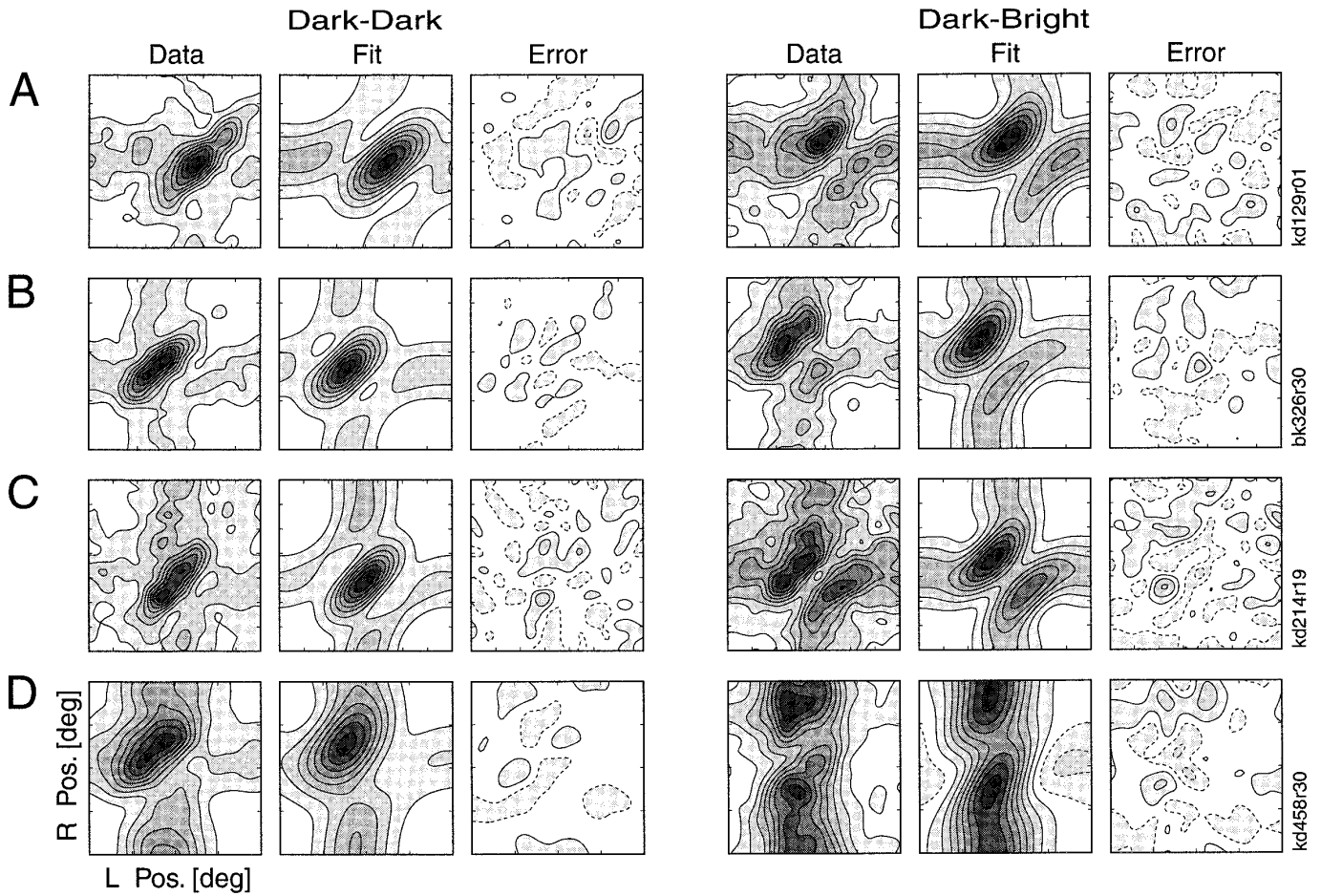


FIG. 12. Four examples (A–D) are shown of complex cell responses that are fit well by a single energy unit model. Data, fit and error plots are presented for each cell. Fits are not decomposed into separate monocular and binocular terms. *Left* 3 panels present responses to dark stimuli for 2 eyes and *right* 3 panels show responses to dark and bright bars for left and right eyes, respectively. X_L - X_R profiles for dark-dark and dark-bright conditions are fit independently. Fit and error plots are shown using the same amplitude scale. Stimulus duration (ms), correlation delay (ms), optimal spatial frequency, orientation, domain size, stimulus size, maximum spikes per stimulus, and laminar location of the cells are A: 52.8, 60, 0.56, 35°, 6 × 6°, 15 × 0.5°, 2.7, 2+3; B: 52.8, 70, 0.4, 167°, 3.5 × 3.5°, 4 × 0.5°, 1.5, 6; C: 39.6, 75, 0.29, 40°, 7 × 7°, 8 × 0.4°, 4.9, 2+3; and D: 52.8, 60, 0.83, 160°, 4 × 4°, 20 × 0.5°, 3.9, 2+3, respectively.

diagonal. For the dark-bright condition (Fig. 13A, *right*), there is a vertical residual contour that cannot be accounted for by the fit. For the cell shown in Fig. 13B, the dark-dark condition provides a very good fit with a nearly perfect even symmetry ($\psi = 0.45^\circ$). However, the dark-bright condition has an unexpected phase value ($\psi = 55^\circ$). For this condition, the energy model predicts an inversion of phase from the dark-dark condition ($\psi = 180.5^\circ$; see Fig. 8, C and D). This latter deviation does not appear to be accounted for by a multiple energy unit model.

To examine how well the disparity-energy model fits data from all complex cells, parameters of the fits to the binocular RF profiles are compared for the four stimulus conditions. This analysis was performed for 40 cells. The disparity-energy model predicts that the amplitude, A_B , of the binocular component in Eq. 8 is the same for the four conditions of bright and dark stimulus combination for the two eyes. It also predicts that the phase, ψ , is the same for the matched polarity conditions, but is different by 180° between matched and opposite polarity conditions. In Fig. 14A, the amplitude

and the phase of the fit to the dark-dark profile are plotted for each cell on a polar coordinate system relative to those of the bright-bright profile. The distance from the origin to a point depicts the amplitude (on a logarithmic scale), whereas the angle represents the phase. The amplitude and the phase of the corresponding bright-bright condition is normalized to the point (A_B, ψ) = (1, 0). Therefore a given point will fall near the coordinate (1, 0) if the dark-dark profile is closely similar to that of the bright-bright condition, as predicted by the disparity-energy model. A dense cluster of points around the expected region demonstrates that the prediction is fulfilled for most complex cells. Of 40 cells, the phase for the dark-dark condition for 36 cases is within 45° of that for the bright-bright condition. The mean \pm SD of the phase is $-3.3 \pm 38.1^\circ$. The relative amplitude of the dark-dark condition with respect to the bright-bright condition is 1.5 ± 1.4 .

Figure 14B presents results of the same analysis performed for mismatched polarity conditions, i.e., bright-dark (\square) and dark-bright (\blacksquare). Again, the amplitude and phase

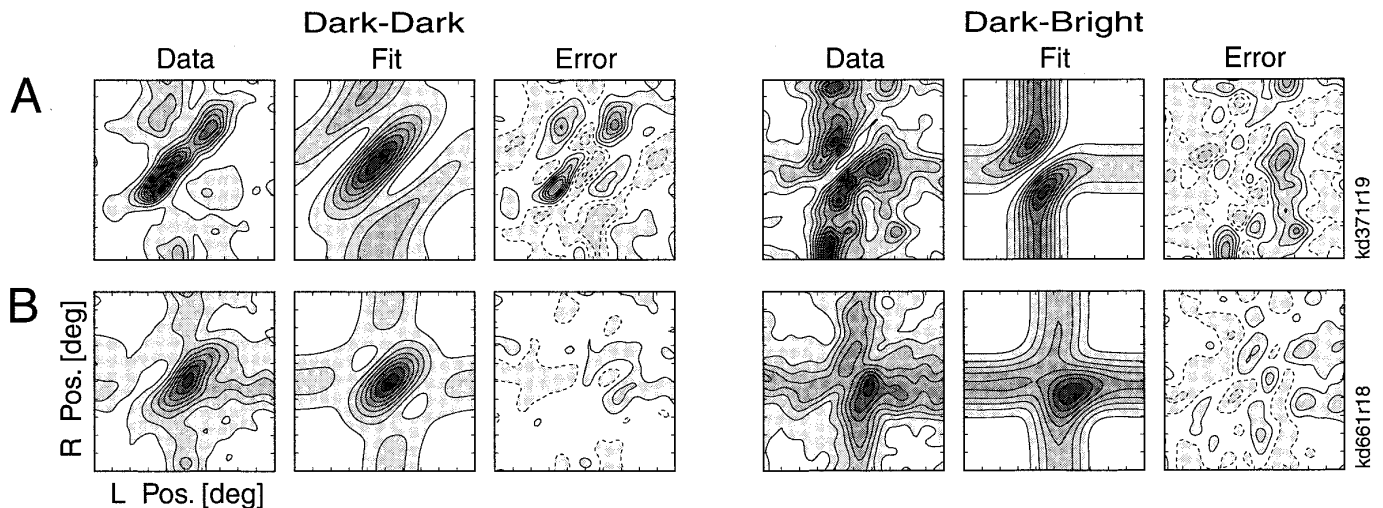


FIG. 13. Two cases (A and B) are shown for which a single energy-unit model did not give satisfactory fits. Results are presented in the same format as that for Fig. 12. Stimulus duration (ms), correlation delay (ms), optimal spatial frequency, orientation, domain size, stimulus size, maximum spikes per stimulus, and laminar location of the cells are A: 52.8, 75, 0.9, 77°, $5.5 \times 5.5^\circ$, $6 \times 0.4^\circ$, 0.6, 2+3; B: 26.4, 50, 0.41, 125°, $10 \times 10^\circ$, $10 \times 0.5^\circ$, 5.2, 6, respectively.

are plotted relative to those of the bright-bright condition. Although the degree of scatter in phase is larger for these conditions, nearly one-half of the points fall near the expected value of $180 \pm 45^\circ$. There is a notable lack of points in the range $0 \pm 45^\circ$. When we examine cells by using multiple criteria for different stimulus conditions, phase was within the expected range for 30 of 40 cells (75%) jointly for one opposite contrast condition and the dark-dark condition. Under the strictest criteria, phase was within $\pm 45^\circ$ of the expected value for all three conditions (dark-dark, bright-dark, and dark-bright) for 13 of 40 cells (33%). Note that this represents a high degree of organization. If phase relationships were completely random, we would expect only 1.6% (0.25^3) of the cells, or less than one cell out of our sample, to satisfy all of the three criteria, because the probability of each phase falling within 45° of a given value is $1/4$. The means \pm SD of phases for bright-dark and dark-bright conditions are $198 \pm 64^\circ$ and $190 \pm 68^\circ$, respectively. On average, amplitude for the mismatched polarity conditions appears to be smaller than that for the bright-bright condition. The means \pm SD are 0.76 ± 0.58 and 0.82 ± 0.63 for bright-dark and dark-bright conditions, respectively. The source of this response amplitude difference between matched and opposite polarity conditions is not clear. Taken together, the disparity-energy model appears to provide a reasonable description of the data for a substantial fraction of complex cells. However, under strict joint criteria, deviations from the predictions of the models are present for the majority of cells.

Quantitative analysis of disparity tuning properties

To analyze the data further without relying on assumptions specific to the single energy unit model, we construct a composite RF profile from the data for all four contrast conditions: BB (bright-bright), DD (dark-dark), BD (bright-dark), and DB (dark-bright). For each cell, the composite binocular RF profile is given by $BB + DD - BD - DB$. Regardless of the exact shape of the binocular response

profile, and therefore without assuming a specific model, this computation cancels monocular contributions to the response and yields a purely binocular response profile. This produces a single composite profile for each cell without discarding any data and incorporates data that are otherwise difficult to deal with, e.g., the lack of binocular response component shown in Fig. 11D and the imbalance of response strengths for bright and dark stimuli (Fig. 12D). This computation is equivalent to a procedure for deriving a second-order Wiener-like kernel in nonlinear systems analysis (Emerson et al. 1987, 1989). For a special case of the single energy unit model, it may be shown that the sum ($BB + DD - BD - DB$) represents the binocular response component isolated in the third term of Eq. 7 and shown in the *rightmost panels* of Fig. 10, A and B. Because it is in this profile that the most interesting information is contained regarding the binocular processing, the single X_L - X_R profile is further processed in accordance with the procedure of Fig. 4 to obtain a disparity tuning curve and a disparity-time (D-T) plot.

Representative data summarized this way from 10 cells are presented in Fig. 15. For each cell (A–J), the *top panel* shows the composite X_L - X_R profile. The *bottom panels* depict the D-T profile and a disparity tuning curve taken at a time delay indicated by a horizontal dashed line. All of the examples show clear diagonal structure in the X_L - X_R profiles. The cells (A–J) are ordered roughly according to the degree of asymmetry in the disparity tuning curves from highly even-symmetric to those that are not. For example, Fig. 15A depicts a cell that had nearly exact even-symmetry, whereas the disparity tuning curves for the cells of Fig. 15, E–I, clearly do not. The asymmetry is obvious in the X_L - X_R profiles as well as in the D-T profiles and disparity tuning curves. The cell shown in Fig. 15J had an inverted disparity tuning curve with central suppressive flank and an excitatory flank on each side. This type of cell is found rarely, and the S/N ratio was relatively low [$(S + N)/N = 12.2$ dB]. The S/N ratio is defined as the energy (sum of squares) of the disparity tuning curve (dots) to the energy of the profile at

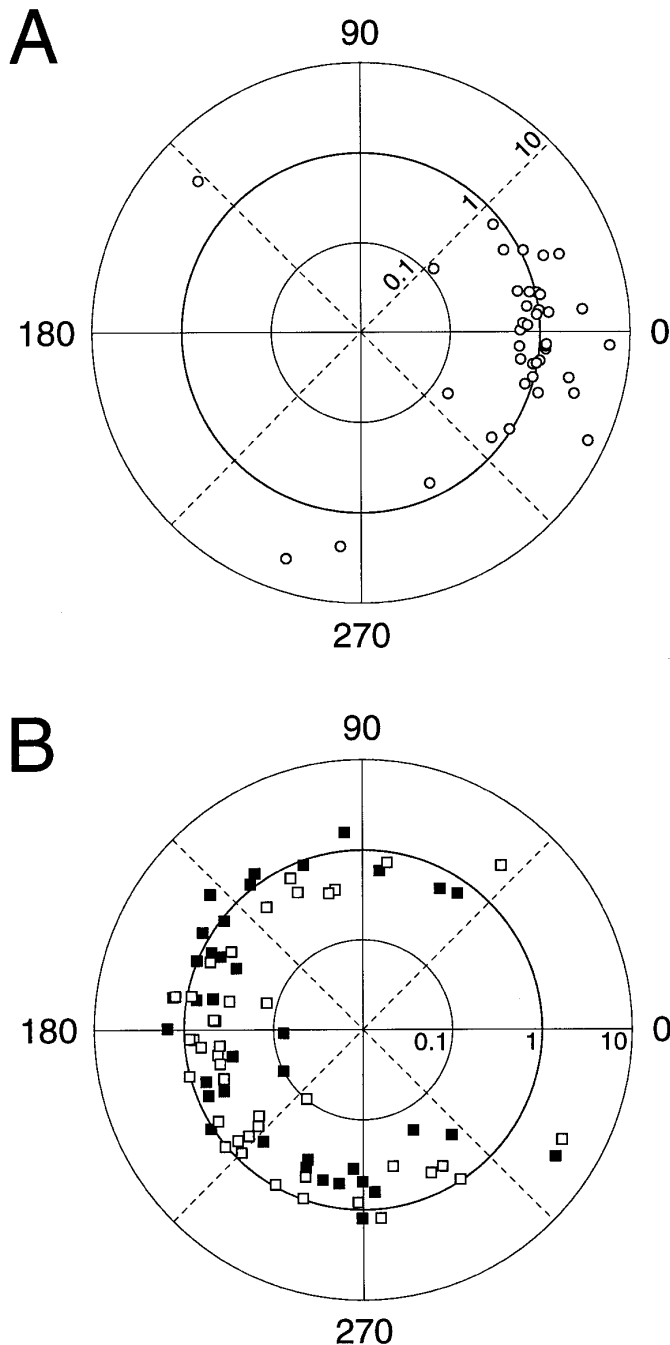


FIG. 14. Consistency of binocular RFs for 4 permutations of sign of stimulus contrast is examined. *A*: amplitude and phase of fit to dark-dark profile are shown in a polar coordinate system. Amplitude and phase are plotted as distance of point from origin (on a logarithmic scale) and angle from right-ward horizontal axis, respectively. Parameter values are relative to those for bright-bright profile. Thus if dark-dark profile is identical to that for bright-bright condition (as predicted by disparity-energy model), point will fall at amplitude = 1 and phase = 0° . Dashed lines indicate $\pm 45^\circ$ and $180 \pm 45^\circ$. *B*: similar plots are shown for fits to binocular RFs for opposite contrast conditions. \square and \blacksquare , relative amplitude and phase for bright-dark and dark-bright conditions, respectively. Again, parameters are shown relative to those for bright-bright condition. Disparity-energy model predicts that points should fall at amplitude = 1, and phase = 180° .

a delay of -50 ms (not shown). Using a negative delay measures correlation with future events. Therefore, by definition, the resulting profile should represent the noise level (Ohzawa et al. 1996). The ratios are given in decibels, $\text{dB} = 10 \cdot \log_{10}(S/N)$. Despite the low S/N ratio for the cell shown in Fig. 15*J*, the mean \pm SE of the phase of the best-fitting disparity tuning curve (obtained by a Levenberg-Marquardt optimization on the X_L - X_R profile) was 148 ± 10.3 degrees. For other cells, the S/N ratios were higher and the SEs for the phase were typically only several degrees.

The results of Fig. 15 illustrate the following points. First, some of the X_L - X_R profiles show highly elongated diagonal regions of excitation as exemplified by the cell of Fig. 15*F* and also by *B* and *D*. For these cells, the extent of the X_L - X_R profiles is much larger along the positive ($+45^\circ$) diagonal than along the negative (-45°) diagonal. Considering that the $+45^\circ$ lines represent constant-disparity lines and frontoparallel planes (Figs. 1 and 4), these cells have binocular RFs that are very wide across the spatial direction (X - Y) but are quite narrow in the depth dimension. For the excitatory subregion of the cell of Fig. 15*F*, its width along the disparity dimension is only 1.7° whereas it covers nearly 6° of space in the frontoparallel plane. Therefore, this cell is nearly four times more selective to changes in disparity than to changes in position along the frontoparallel plane.

Second, none of the D-T profiles show any appreciable tilt of subregions in the disparity-time domain. This was noted for the cells of Figs. 5 and 6, and we have found this to be the case for nearly all of our cells. The result indicates that these cells do not respond selectively to changes in disparity over time, such as those that occur during motion-in-depth. This may be well understood by an analogy to the relationship that exists between the tilt of monocular space-time (X - T) RFs and cells' velocity selectivity. It has been shown for simple cells that the slope of X - T RFs predicts the preferred velocity and direction of the neuron (DeAngelis et al. 1993a; McLean and Palmer 1989; McLean et al. 1994). Just as a simple cell that shows no tilt in its X - T RF does not exhibit direction selectivity, a lack of tilt in the D-T profile of a complex cell implies that the cell will not exhibit any direction preference for motion-in-depth. Considering that the majority of simple cells exhibit some degree of space-time inseparability (DeAngelis et al. 1993a,b; McLean et al. 1994), the lack of tilt of D-T RFs is striking. We also note that the temporal responses seen in these D-T plots are monophasic, i.e., they are initially either positive or negative, reach a single peak, and return to zero without an inversion of the sign of the response. This is in contrast to the temporal responses of simple cells for which multiphasic temporal impulse responses are the rule (DeAngelis et al. 1993a,b). These points are examined quantitatively below.

We have analyzed further the disparity tuning of complex cells by fitting a Gabor function to the disparity tuning curves as shown in the *bottom panels* of Fig. 15, *A*-*J*. The use of a Gabor function here is primarily for the convenience of extracting intuitive parameters from the data. For a single energy unit model based on Gabor subunits (see Eq. 7), it may be shown mathematically that the disparity tuning curve (to a thin bar stimulus) is a Gabor function because the

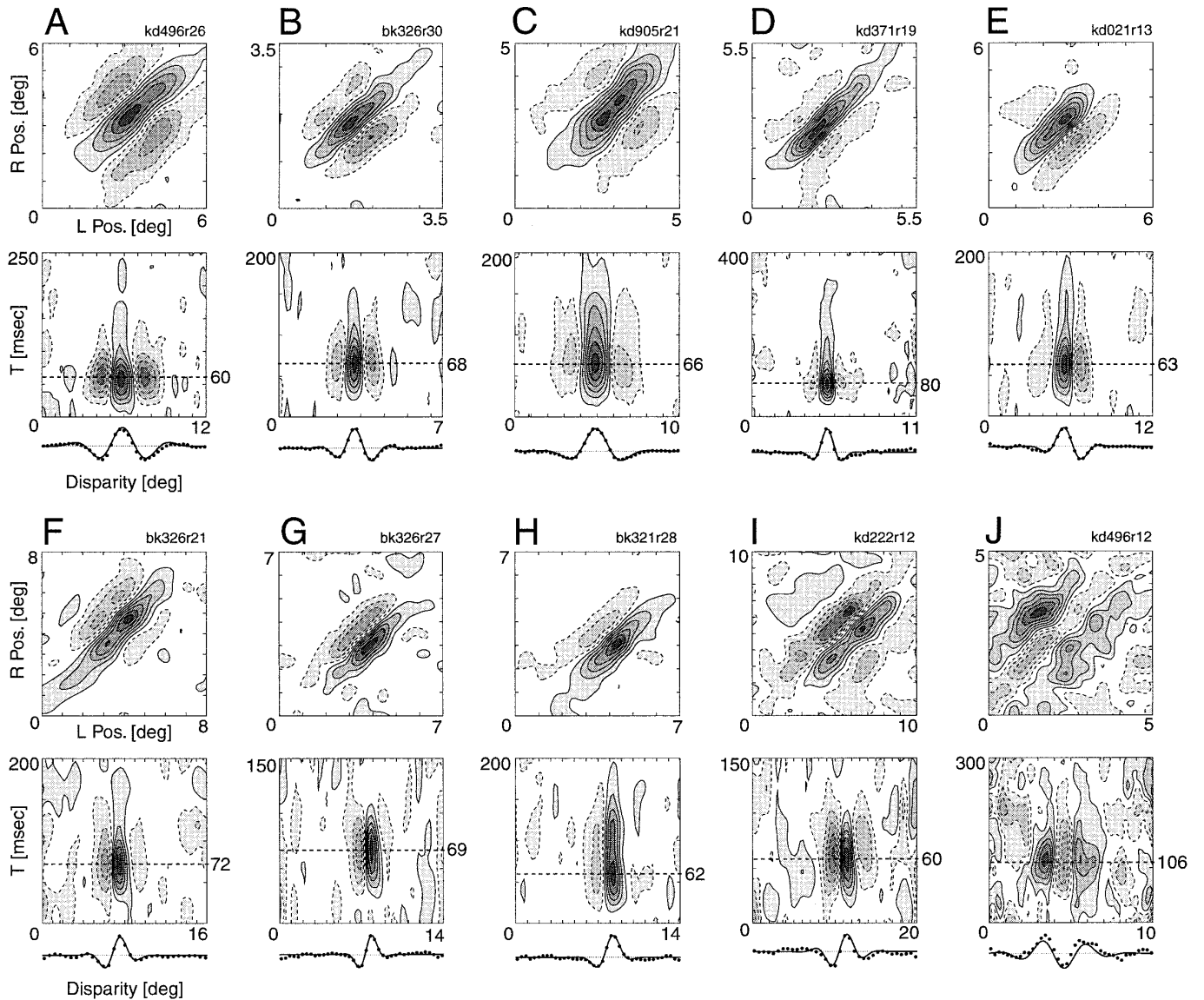


FIG. 15. Composite binocular RFs and disparity tuning curve are shown for 10 complex cells (A–J) in X_L – X_R and disparity-time (D–T) domains. Composite profile is computed as a linear combination of responses to 4 permutations of sign of contrast for 2 eyes (sum of matched contrast responses minus sum of opposite contrast responses). This computation cancels the monocular response terms. — (in each D–T plot), optimal correlation delay for which the disparity tuning curve is shown below. Dots in disparity-tuning curves represent raw data to which a Gabor function (5 free parameters) is fit. S/N ratios (see text) in decibels (dB) for cells are A, 21.4; B, 16.7; C, 20.5; D, 16.7; E, 16.2; F, 12.9; G, 10.1; H, 15.5; I, 11.8; and J, 12.2. Laminar locations are A, 2+3; B, 6; C, 2+3; D, 6; E, 6; F, 2+3; G, 6; H, 2+3; I, 6; and J, 6. Stimulus duration (ms), optimal spatial frequency, orientation, stimulus size, and maximum spikes per stimulus for cells are E, 52.8, 0.5, 45° , $20 \times 0.3^\circ$, 3.5; G, 52.8, 0.77, 155° , $5 \times 0.5^\circ$, 2.4; H, 52.8, 0.49, 20° , $20 \times 0.5^\circ$, 2.1; I, 39.6, 0.14, 1° , $20 \times 0.5^\circ$, 1.4; J, 52.8, 0.57, 140° , $20 \times 0.5^\circ$, 0.3, respectively. Details for other cells are presented in previous figure legends.

integral of a two-dimensional Gabor function along the length of subregions is a Gabor function (Ohzawa et al. 1990). However, the function provides reasonable fits to the disparity tuning curve of most of the cells, even for those whose X_L – X_R profiles cannot fully be account for by the energy model. The fitting procedure is performed using a modified simplex optimization algorithm (Press et al. 1992). Interactive graphical software is used to set initial parameters of the fit manually, so that the algorithm is less likely to get trapped in a local minimum. The convergence of the algorithm is monitored graphically during fitting. In the disparity tuning plots, data points are shown as filled circles, and

represent a time-slice of the data taken at the correlation delay indicated by a horizontal dashed line in the D–T plot. A function of the following form is used for the fit

$$A \cdot \exp[-k_D(d - C)^2] \cdot \cos[2\pi f_D(d - C) + \psi_D] \quad (9)$$

where A is a scaling factor, d is the independent variable for disparity, and C is the center of the Gaussian envelope of the fitted function. The phase parameter ψ_D of the Gabor function indicates the symmetry of the disparity tuning curve. The parameter k_D determines the width of the Gabor function, and hence the total width of the disparity tuning curve. The parameter f_D is the disparity frequency, which is

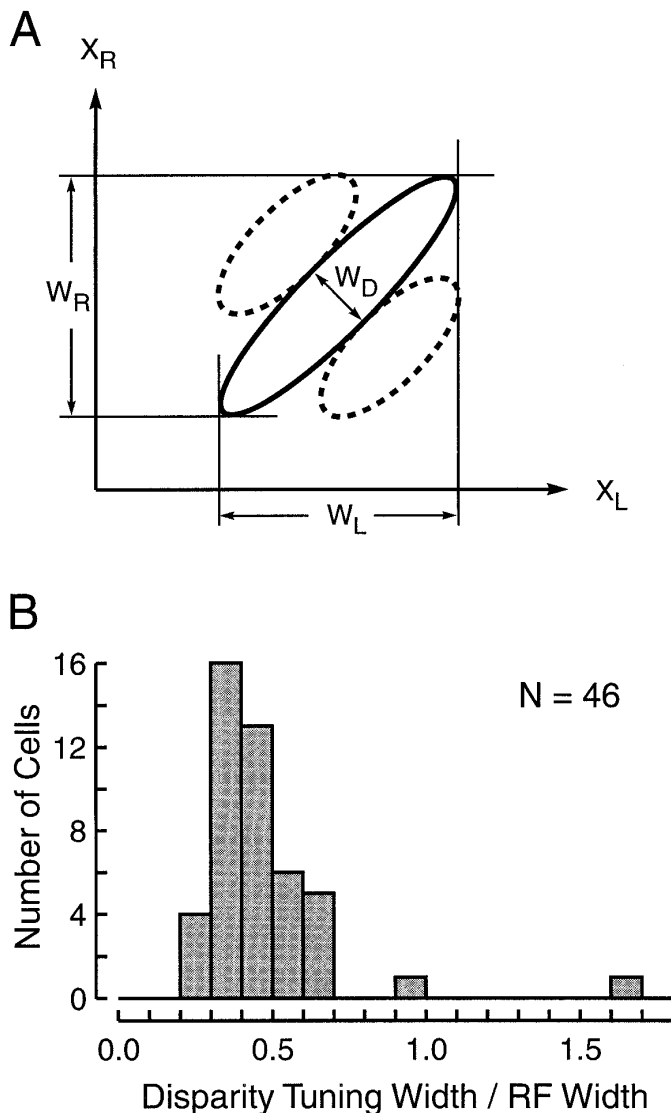


FIG. 16. Width of disparity tuning is compared with size of monocular RFs. *A*: ratio, W_D (width of elongated excitatory area) divided by average of W_L and W_R (left and right RF widths, respectively), is used as a metric of disparity-tuning width relative to RF width. W_D is obtained from disparity-tuning curve as $1/2$ of inverse of disparity frequency, because X_L - X_R domain is distorted geometrically. W_L and W_R are estimated from extent of contour plots, typically from a contour level that is 5 or 15% of peak. If contour at this level is not contained in X_L - X_R domain, a higher level is used by incrementally adding 10%. Highest level used was 55% of peak. For any level, L_C , used for measurement of RF width, a correction is made to W_D by factor $(\cos^{-1} L_C)/(\pi/2)$. *B*: histogram is shown of ratio of W_D to RF width.

defined as the frequency at which disparity tuning curve alternates between excitation and suppression. Together, k_D , and f_D determine the width of an excitatory (or suppressive) region. Therefore these parameters determine how narrowly a cell is tuned to changes in disparity. Relationships between these disparity tuning parameters and monocular RF parameters are described below.

Relation between disparity tuning and RF width

As we noted above for Fig. 15*F*, the selectivity of a neuron in the disparity dimension can be much sharper than

its selectivity for position. We now quantitatively evaluate how narrow the disparity tuning is with respect to the width of the overall RF. As shown in Fig. 16*A*, we derive an index of disparity tuning width relative to RF width as a ratio

$$W_D / [(W_L + W_R)/2] \quad (10)$$

where W_D is the width of an excitatory region in the disparity dimension, and W_L and W_R are the RF widths for the left and right eyes, respectively. W_D is obtained from the Gabor fit to the disparity tuning curve and is given by $1/(2f_D)$, which corresponds to half of the period of the sinusoid at the disparity frequency. W_L and W_R are measured directly from contour plots in the X_L - X_R domain by the strongest excitatory region. Figure 16*B* shows the distribution of the ratios for our sample. The sample size here is slightly larger than that for Fig. 14 ($n = 40$), because the composite profiles used here provide more reliable fits and estimation of parameters than the raw profiles used for Fig. 14. The mean \pm SD of the ratios is 0.46 ± 0.21 . This indicates that the majority of cells are more narrowly selective to binocular disparity than to stimulus position, by a factor of two or more.

Relationships between disparity tuning and monocular properties

A summary of the relationship between the phase ψ_D of the disparity tuning curve and the cell's preferred orientation is shown in Fig. 17*A*. This is of interest, because the single energy-unit model predicts a tight relationship between ψ_D and the phase difference, ψ , between the left and right RFs of subunits (see Fig. 10*B* and Eq. 7). In addition, we have shown for simple cells that there is an anisotropy in the distribution of the phase difference between left and right RFs. Specifically, cells tuned to near horizontal orientations tend to have similar RFs (hence small ψ) for the two eyes, whereas those tuned to oblique or vertical orientations have a variety of phase differences (DeAngelis et al. 1991, 1995a; Ohzawa et al. 1996). This anisotropy is consistent with the hypothesis that the visual system employs an optimized encoding scheme that takes advantage of a statistical bias in the distribution of horizontal versus vertical disparities, which results from lateral placement of the two eyes. Although the single energy unit model is not always sufficient for some cells, as we have seen above, we still may find a relationship between the preferred orientation and ψ_D if these simple cells are the subunits that feed into the complex cells.

For our sample of complex cells, there is no obvious trend in the distribution shown in Fig. 17*A*. The data points do not form a triangle-shaped distribution with vertices at bottom left, top right, and bottom right corners similar to the distribution for simple cells (DeAngelis et al. 1991; Ohzawa et al. 1996). In particular, there are cells tuned to near horizontal orientations (0) that exhibit a substantial asymmetry in the disparity tuning curve. Moreover, most cells have a ψ_D value that is <90 . There is no statistically significant dependence of phase of the disparity tuning curve on preferred orientation (linear regression analysis; $P = 0.4$).

We also have examined the relationship between the distribution of ψ_D and cells' preferred spatial frequency. This distribution, shown in Fig. 17*B*, appears uniform with cells having a wide range of symmetry at every spatial frequency.

The sample size is smaller by two cells than that for Fig. 17A, because the spatial frequency tuning was not measured for these cells. Again, there is no statistically significant dependence of ψ_D on spatial frequency (linear regression analysis in the phase vs. log spatial frequency domain; $P = 0.26$). Considering that simple cells do not show any dependence on preferred spatial frequency with regard to the phase difference between left and right RFs (DeAngelis et al. 1995a; Ohzawa et al. 1996), it is not surprising to find a relatively uniform distribution for complex cells.

We also have examined a possible relationship between the disparity frequency, f_D , and the optimal spatial frequency (measured by monocularly presented sinusoidal drifting gratings). An inspection of Eq. 1, 7, and 9 shows that, for the single energy unit model, the disparity frequency and the optimal frequency of monocular RFs should be the same, i.e., $f_D = f$. It is of interest to find out if this relationship holds for our sample of complex cells. Results are shown in Fig. 17C. It is clear that, for most cells, the disparity frequency is substantially lower than the optimal spatial frequency of the cell measured monocularly (average for the two eyes). An exact match of the frequencies is indicated by the solid oblique line in Fig. 17C. A linear regression analysis (in the log-log domain) shows a significant dependence of disparity frequency on monocular spatial frequency ($P < 0.005$, slope = 0.29, correlation coefficient = 0.42). Note, however, that the slope of the best-fitting line is substantially < 1.0 . The cell shown in Fig. 11, A and B, whose binocular response is fitted well with a single energy unit model, exhibits a relatively similar disparity frequency (0.25 c/deg) and optimal spatial frequency (0.35 c/deg). However, the cell presented in Fig. 11, C and D, had a disparity frequency (0.27 c/deg) less than 1/3 of the optimal spatial frequency (0.93 c/deg) measured monocularly. The bar width of the stimuli used for mapping was 0.4 for this cell, and spatial blurring cannot explain the difference of a factor of 3. Linear regression analysis shows a trend for cells tuned to high spatial frequencies to have high disparity tuning frequencies as well. Although a single energy unit model may be reasonable for those cells that have similar f_D and f , clearly, disparity tuning of other cells is generated through a much more complex mechanism. Sources of this deviation are unknown. Note that they are not due to a geometric distortion of the X_L - X_R domain as described above, because the distortion is already factored into the transformation from the X_L - X_R space into the disparity domain (Fig. 4). This deviation from the energy model prediction is considered further in DISCUSSION.

Another parameter we have examined is the relationship between the optimal spatial frequency and the disparity range, which represents the extent of the disparity tuning curve. Results of this analysis are shown in Fig. 17D. The disparity range is defined as one half of the width of the Gaussian envelope of Eq. 9 at 5% of the peak. This is a measure of the largest disparity offset, from the center of the disparity tuning curve, that may still elicit binocular interactions. Therefore, if the center of the disparity tuning curve is at zero disparity, then the value of the disparity range represents a disparity limit for binocular interactions of a given neuron. The data are of interest in relation to the size-disparity correlation that have been suggested as a key

property of a model (Marr and Poggio 1979). The data are also relevant to related psychophysical evidence (Legge and Gu 1989; Schor and Wood 1983; Smallman and MacLeod 1994). A linear regression analysis (in the log-log domain) shows significant dependence of disparity range on monocular spatial frequency ($P < 0.005$, slope = -0.39 , correlation coefficient = -0.52). The trend indicates that neurons tuned to higher spatial frequencies are likely to have smaller disparity ranges. However, the disparity range shrinks with spatial frequency at a rate that is substantially less than the exact inverse relationship (slope = -1).

We also have examined the number of subregions in the X_L - X_R map, i.e., the number of positive and negative peaks in the disparity tuning curve (see Fig. 15). Because the disparity range shrinks with spatial frequency at a rate less than -1 (Fig. 17D), there may be more peaks in the disparity tuning curve for cells tuned to high spatial frequencies than those tuned to low frequencies. Consequently, there may be a greater degree of ambiguity in disparities signalled by cells tuned to higher spatial frequencies, because the preferred disparity is no longer unique for these cells. Results shown in Fig. 17E indicate that this is not the case. Although there is some scatter, for the majority of cells, the number of disparity subregions is between three and four [3.44 ± 1.07 , (mean \pm SD)], and there is no significant dependence on spatial frequency ($P = 0.38$). This is primarily due to the fact that the trends shown in Fig. 17, C and D, cancel each other. Therefore there is no greater degree of ambiguity in disparities signaled by cells tuned to high spatial frequencies. In this sense, the disparity tuning of complex cells does not suffer from an analogous monocular ambiguity problem presented by simple cell RFs, where there tend to be more RF subregions for cells tuned to high spatial frequencies (see Fig. 13a of DeAngelis et al. 1995a).

Sensitivity to motion-in-depth

As we have noted above, the data of Fig. 15 show that nearly all binocular RF profiles are oriented vertically in the D-T domain. Qualitatively, this predicts that these neurons are not sensitive to motion-in-depth (Cynader and Regan 1978, 1982; Spileers et al. 1990). We now examine this issue quantitatively.

D-T profiles show how a neuron's preferred disparity changes over the time course of the response. One way to determine the rate of change in preferred disparity is to fit a straight line to the subregions and determine its slope. The slope, $\Delta\text{disparity}/\Delta\text{time}$, gives the rate of change of preferred disparity, which may be defined as the preferred velocity-in-depth. However, there is a practical problem with this procedure because there are typically multiple subregions in D-T profiles as shown in Figs. 18A and Fig. 15. A better estimate of preferred velocity-in-depth may be obtained by a frequency domain analysis, using an analogous procedure to that devised for determining the preferred velocity of simple cells from their space-time (X-T) receptive field (DeAngelis et al. 1993a; Ohzawa et al. 1996). First, a two-dimensional Fourier transform (Bracewell 1978; Press et al. 1992) of the DT profile is computed. Figure 18B presents the amplitude spectrum computed from the transform of the DT profile in Fig. 18A. The data of Fig. 18B

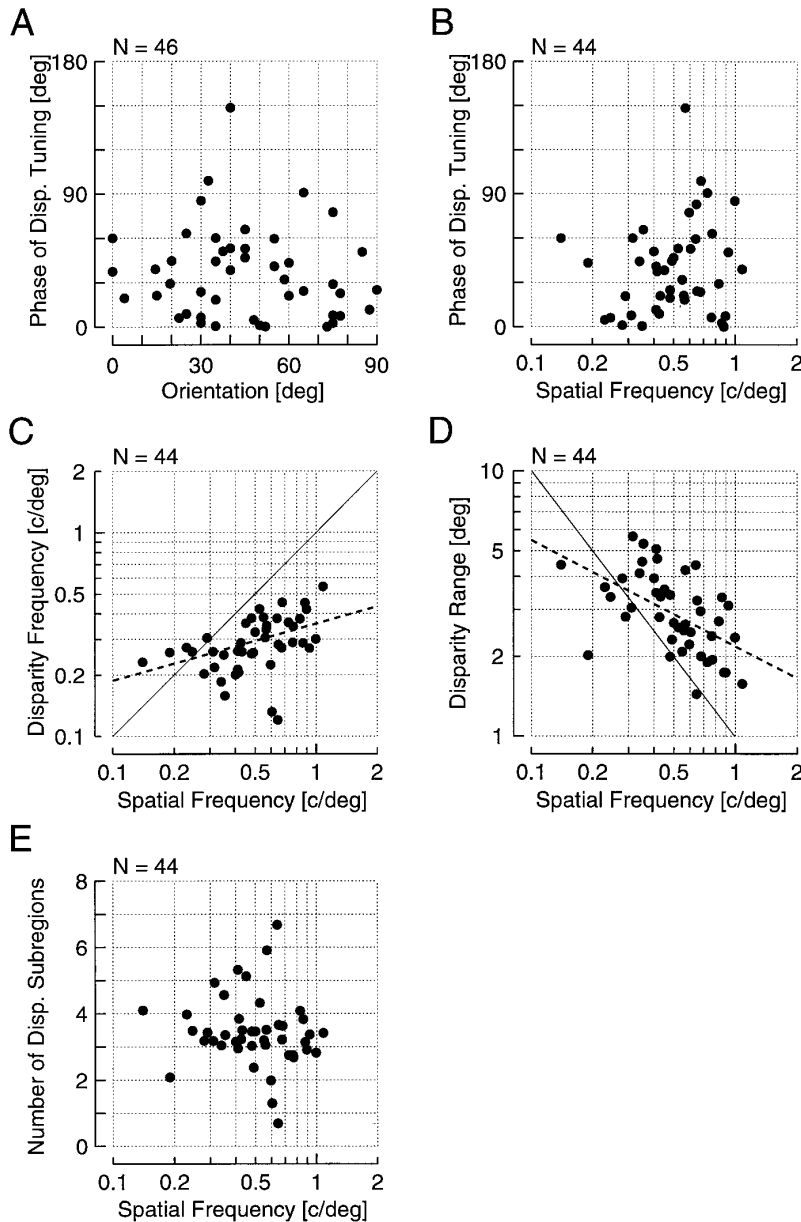


FIG. 17. *A*: phase of disparity-tuning curve is plotted against preferred orientation for each cell. When preferred orientations are different for the 2 eyes, an average value is used. Original phase domain of 0–360° is folded into 0–180° such that 0 and 90° represent even- and odd-symmetric disparity-tuning curves, respectively. Similarly, original orientation domain 0–180° is folded into 0–90° such that 0 and 90° represent horizontal and vertical, respectively. *B*: phase is plotted against monocular preferred spatial frequency. Again, average of left and right spatial frequencies is used. Monocular preferred spatial frequency is measured using drifting sinusoidal gratings of optimal orientation, 40% contrast, and a temporal frequency of 2 Hz. Sample size is smaller by 2 than that for *A* because spatial frequency tuning curves were not recorded for 2 cells. *C*: disparity frequency is plotted against preferred spatial frequency for each cell. Energy model predicts a match (indicated by a solid diagonal line) between these 2 frequencies. — — —, linear regression fit in log-log domain. *D*: disparity range, which is a metric of extent of disparity tuning curve, is plotted against monocular preferred spatial frequency. —, inverse relationship between disparity range and spatial frequency; — — —, linear regression fit. Value is computed from parameter k of Eq. 9 as $\sqrt{\{-\ln(0.05)\}/k} = 1.73/\sqrt{k}$, where $\sqrt{}$ and \ln represent square root and natural log functions, respectively. *E*: number of subregions in disparity tuning curve is plotted against monocular spatial frequency. It is defined by $4 \cdot (\text{disparity frequency}) \cdot (\text{disparity range})$ and derived from values plotted in *C* and *D*.

represent a tuning surface in the disparity frequency-temporal frequency domain. From the location of the peaks, the preferred disparity frequency, f_{Dopt} , and temporal frequency, f_{Topt} , are determined. The preferred velocity-in-depth, V_{Dopt} , is given by

$$V_{\text{Dopt}} = f_{\text{Topt}}/f_{\text{Dopt}} \quad (11)$$

For the cell presented in Fig. 18, V_{Dopt} is 2.8 deg/s. For comparison, the preferred velocity of this cell for monocular motion (and presumably binocular motion along the front-parallel plane) is 18.4 deg/s, as estimated by the ratio of optimal temporal frequency to spatial frequency determined by separate measurements using drifting sinusoidal gratings (Baker 1990). In other words, the preferred velocity-in-depth for this cell is more than six times slower than the preferred velocity for monocular motion. Unfortunately, we cannot perform this comparison for all of our sample of complex cells because we did not measure temporal fre-

quency tuning for other cells. However, comparisons of preferred velocity distributions suggest that this is a general finding. Figure 19*A* presents a histogram of preferred velocity-in-depth for our sample of complex cells. Most of the complex cells prefer slow velocity-in-depth, typically <4 deg/s. The histogram is shown with logarithmically scaled bins (Movshon 1975), because the velocity values span a large range and differences near zero velocity are important. In contrast, Fig. 19*B* shows a monocular preferred velocity distribution for simple cells. Most of these cells prefer monocular velocities >4 deg/s. The simple cell data are replotted with logarithmically scaled bins from Fig. 17 of DeAngelis et al. (1993a). The difference between the two distributions in Fig. 19, *A* and *B*, is statistically significant ($P < 0.005$, t -test and Kolmogorov-Smirnov test). An equivalent histogram is not available for complex cells because monocular mapping of receptive fields reveals only the envelope and no subunit structure that determines their spatial and temporal

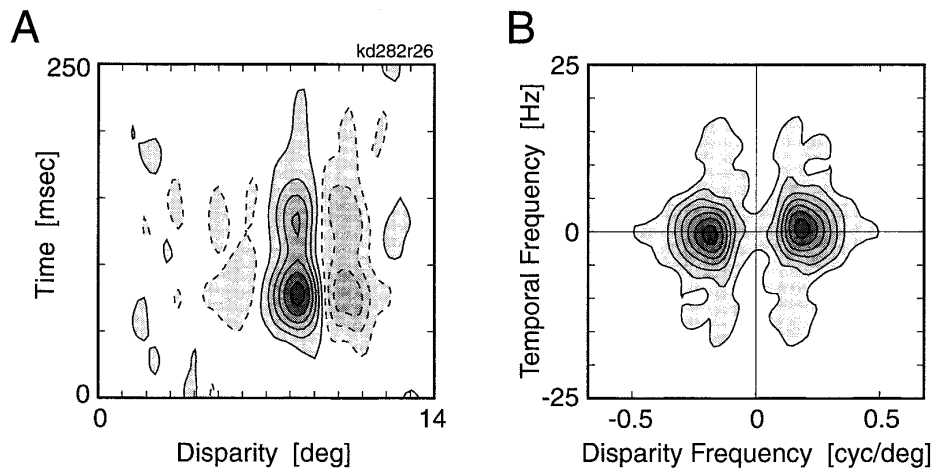


FIG. 18. A frequency domain analysis is performed on a disparity-time RF of a complex cell. *A*: disparity-time plot of composite RF of a complex cell shown in a format similar to that of Fig. 15. Stimulus duration (ms), correlation delay (ms), optimal spatial frequency, orientation, stimulus size, maximum spikes per stimulus, and laminar location of cell are: 39.6, 75, 0.34, 35°, $10 \times 0.4^\circ$, 2.1, 6, respectively. *B*: amplitude spectrum shown here is obtained by a 2-dimensional fast Fourier transform of data in *A*. By locating peak in amplitude spectrum (2 peaks are same and symmetric with respect to origin), optimal temporal frequency and disparity frequency are determined. Preferred velocity-in-depth is estimated by ratio of these frequencies. For all cells, disparity frequency estimated by this method matches closely value obtained by a Gabor function fit to disparity-tuning curve as described for Fig. 15. A 2-dimensional cubic spline is used to locate optimal disparity frequency in amplitude spectrum. Optimal temporal frequency is obtained by fitting a Gaussian to vertical cross-section of spectrum at optimal disparity frequency. Note that optimal temporal frequency here is not the same as that determined by monocular drifting sinusoidal grating stimuli. The former is related to a change in preferred disparity over time, whereas the latter is primarily determined by temporal frequency tuning of subunits (Gaska et al. 1994).

selectivities (DeAngelis et al. 1995b; Ohzawa et al. 1990). As the next best option, Fig. 19C presents the preferred velocity data from complex cells from a previous study (Movshon 1975). The distribution of Fig. 19C appears to be shifted slightly to higher velocities than our sample of simple cells shown in Fig. 19B. Our simple cell distribution shown in Fig. 19B also differs substantially from the simple cell data of Movshon (1975). However, this may be due to a difference in the range of eccentricities from which cells were sampled (DeAngelis et al. 1993a). Taken together, there is little change in the preferred disparity over the time course of complex cell response, as illustrated in the DT profiles of Figs. 15 and 18A. This is confirmed by the quantitative analysis of preferred velocity-in-depth.

Origin of opposite contrast responses

For the complex cells presented in Figs. 3A and 5, clear responses are evident to dichoptic stimuli that are opposite in sign of contrast for the two eyes. For these cases, the responses occur at two disparities, one crossed and the other uncrossed with respect to the preferred disparity for same-contrast stimuli. There are two possible explanations as to why this might occur. We have examined these two possibilities in a control experiment. One explanation is based on the energy model (see Fig. 8D and Eq. 6). A more intuitive account is illustrated in Fig. 20, A and B. For the subunit (S1) at the top of Fig. 8A, redrawn here as Fig. 20A, it is clear that bright bars presented to the two eyes at the center of the subunit RFs elicit the maximal response from this subunit. If we invert the sign of stimulus contrast for the right eye (thus introducing a contrast sign mismatch), this has the equivalent effect of inverting the RF profile for this

eye because of the approximate linearity of the subunit RF, as shown in Fig. 20B (Ohzawa and Freeman 1986b; Ohzawa et al. 1990). For this condition, maximum responses are obtained for two combinations of the bar positions, because there are two equal dark-excitatory regions for the right eye. For other subunits of the energy model, similar displacements of the preferred disparity occur.

The other possible explanation is based on a requirement for the matching of stimulus edges. This is illustrated in Fig. 20, C and D. For the matched contrast condition, both the left and right edges of the stimuli match at only one disparity (Fig. 20C). For the opposite contrast condition, as shown in Fig. 20D, there can only be a partial match of the edges: one with the combination of L and R-1, where the rising edge of the bright left stimulus is matched with the rising edge of the right dark stimulus. The other edge of the dark right stimulus is not matched. The other match is obtained with L and R-2, where the trailing edges are matched.

These two possibilities predict widely different results for variations of the width of the bar stimuli. For the former (energy model) hypothesis, the disparity separation of the two excitatory bands for opposite contrast conditions should be approximately equal to one-half of the period of the subunit RF and should not be highly sensitive to variations in bar width. For the latter explanation, based on edge-polarity matching, the disparity separation should be equal to twice the bar width as shown in Fig. 20D. And the disparity offset (from the optimal disparity) of the excitatory bands for the opposite contrast condition should be equal to the bar width.

Results of a control experiment are shown in Fig. 21. X_L - X_R profiles are shown for bar widths of 0.2, 0.4, and 0.8° , and the corresponding disparity tuning curves (see Fig. 4) are shown below. The distance between the peak of the central

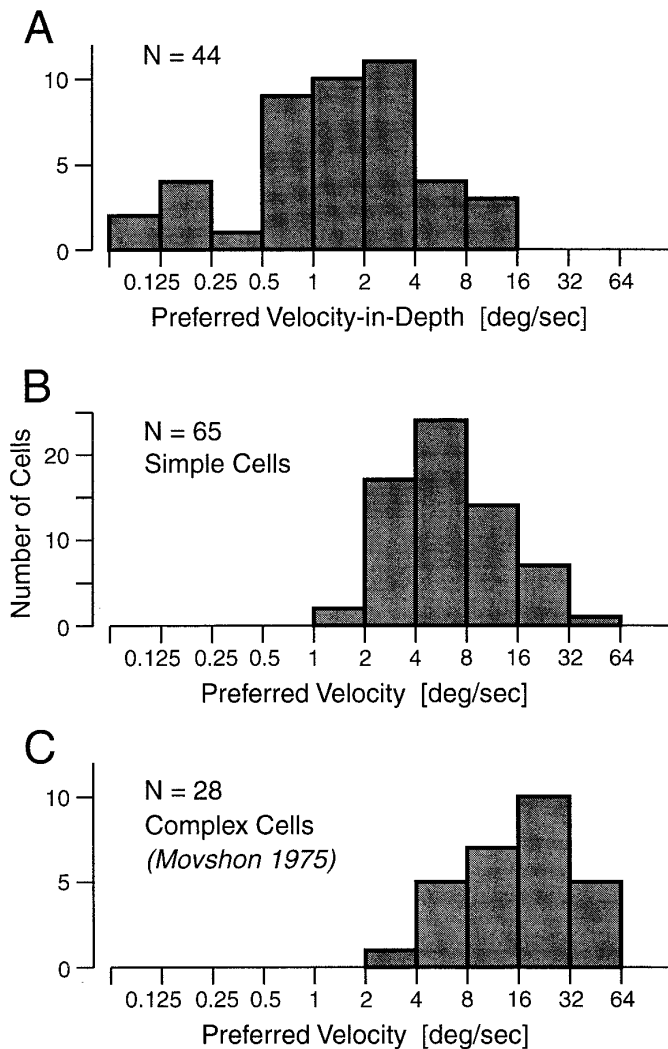


FIG. 19. Distribution of preferred velocity-in-depth is presented and compared with monocular velocity distributions of cortical neurons. *A*: histogram of preferred velocity-in-depth is shown with logarithmically scaled (octave-wide) bins. *B*: distribution of monocular preferred velocities is given for simple cells recorded from same animals from which complex cells in *A* were sampled. *C*: optimal velocity data for complex cells from Movshon (1975) are presented in same format.

excitatory band (solid contours) and that of the suppressive bands (dashed contours) remains constant at $\sim 1.5^\circ$ for all three stimulus bar widths, and therefore no dependence on bar width is observed. On the other hand, optimal subunit spatial frequency is estimated to be 0.29 c/deg by a spatial frequency tuning measurement using drifting sinusoidal gratings (Movshon et al. 1978a). This agrees well with the disparity frequencies obtained by fitting of the one-dimensional disparity tuning curves: 0.24, 0.30, and 0.25 c/deg for the bar widths 0.2, 0.4, and 0.8°, respectively. These results are in agreement with the interpretation shown in Fig. 20, *A* and *B*, and therefore strongly suggest that the origin of responses at nonoptimal disparities under opposite contrast conditions is the multiple subregions of the subunit RFs.

DISCUSSION

We show here that complex cells respond in a characteristic manner that allows a high degree of sensitivity to changes

in binocular disparity while achieving invariance to alterations in stimulus position along the frontoparallel plane. We present a disparity energy model for these complex cells. The model employs a hierarchical organization as originally proposed by Hubel and Wiesel (1962), in which the output of multiple simple-cell subunits is combined to produce the RF of a complex cell. The model provides a remarkably good fit to the data from many complex cells, but deviations from the model are found for some neurons. We now consider implications of our results for the general problem of stereopsis. We also examine potential problems in experimental procedures and interpretation of our data.

Are responses to opposite contrast stimuli undesirable or beneficial?

At first glance, the responses at nonoptimal disparities for opposite contrast conditions appear to be an undesirable phenomenon that adds to ambiguity to the problem of stereoscopic matching. From the responses of one neuron, it is not possible to determine whether the stimulus was a pair of bright bars at the optimal disparity or whether it was a pair of opposite contrast bars at another disparity. This ambiguity must be resolved by additional processing (Blake and Wilson 1991). However, as we have seen above, these responses arise from the fact that the subunit RFs possess multiple alternating subregions similar to those of simple cells (Fig. 20*B*). Therefore the opposite-contrast responses are a natural consequence of the underlying Gabor-like RF profiles. For this reason, the kind of ambiguity we observe here is not unique to stereopsis. For example, responses from one simple cell alone are not capable of signaling whether a bright bar is flashed in an ON (bright-excitatory) flank or a dark bar is positioned in an OFF (dark-excitatory) area.

Contrary to the negative implications of the responses to opposite contrast targets, as described above, there are some advantages to this behavior. Although binocular viewing of a single isolated bar stimulus never results in a combination of opposite contrast to the two eyes, these conditions do occur for extended patterned stimuli under normal viewing conditions. Because cortical neurons are tuned to a limited spatial frequency band, we may consider responses of a given neuron by using the band-limited version of visual stimuli prefiltered to the pass band of the cell. Any visual stimulus will elicit essentially the same response as that to a semiperiodic pattern of bright and dark bands that alternate approximately at the cell's optimal spatial frequency (Fleet et al. 1996; Marr 1982; Marr and Poggio 1979). This condition is illustrated schematically in Fig. 22*A* for the X_L - X_R domain. Segments of a dark-bright-dark sequence are shown as stimuli to the left and right eyes along the horizontal and vertical axes, respectively. Consider a complex cell that has an X_L - X_R map as shown (solid contour indicating excitation to matched contrast and dashed contours indicating excitation to opposite contrasts as in Fig. 15). The binocular combination of these stimuli are excitatory for the cell everywhere in the X_L - X_R domain. For example, the central bright portions of the left and right stimuli cause excitation because they fall exactly in the diagonal excitatory band for matched stimuli (○). Dark stimuli on both ends similarly fall on the diagonal excitatory band with matched contrast sign (●).

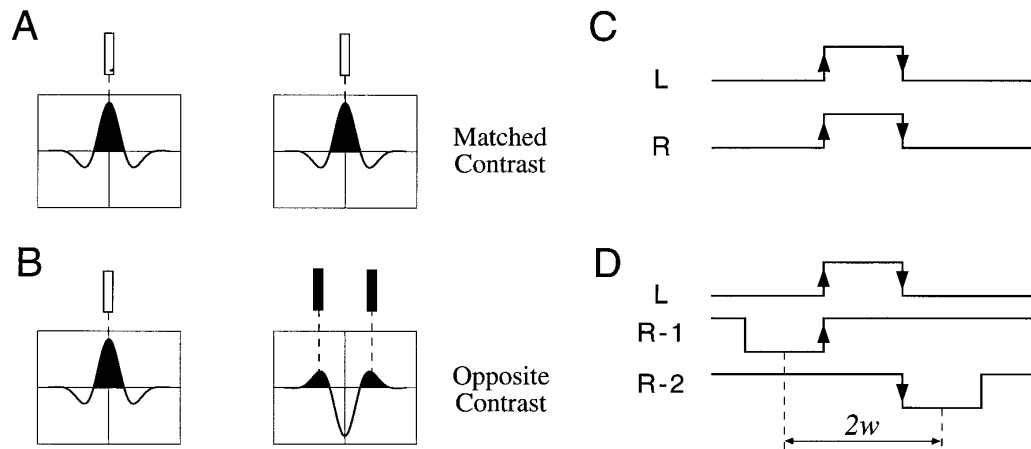


FIG. 20. Two hypotheses are presented to account for origin of responses to opposite contrast stimuli for 2 eyes. *A*: energy model predicts that responses to opposite contrast stimuli originate from presence of multiple subregions in linear subunits. For linear subunits and simple cells, an inversion of sign of stimulus contrast causes apparent inversion of RF profile. *B*: an alternative hypothesis is based on matching of stimulus edges. This hypothesis predicts that excitatory regions to opposite contrast stimuli are separated by twice the stimulus width, w .

However, there are also other combinations of left and right stimulus elements (individual bright and dark bars) as shown by horizontal and vertical dashed lines, many of which are opposite in the sign of contrast for the two eyes. Note that these opposite combinations (bipartite circles) fall exactly on the appropriate regions in the X_L - X_R map (dashed contours), thus providing additional excitation for the neuron. Therefore a periodic binocular stimulus with the appropriate disparity is more effective than a single bar stimulus, even though such stimuli generate a large number of stimulus combinations with contrast sign mismatches.

Figure 22*B* shows that exact locations of the stimuli with respect to the cell's RFs are not important as long as the binocular disparity remains unchanged. Stimuli in this condition are shifted by 90° in phase for the two eyes. Again, it is clear that various combinations of individual bright and dark segments of stimuli fall into appropriate regions of the binocular RF, providing much larger overall excitation than a single contrast-matched target. Therefore, not only is the ambiguity problem due to opposite contrast not unique to complex cells and the disparity encoding problem, these responses are actually beneficial for the visual system in

the natural environment where these mismatches occur in abundance.

Role of complex cells in solving the binocular correspondence problem

The stimulus configurations shown in Fig. 22, *A* and *B*, are similar to that of Fig. 1*D*, which illustrates a large number of possible "false matches" that arise in binocular viewing of natural stimuli. Figure 1*D* originates from Julesz (1968, 1971), and this figure has been duplicated in subsequent articles and books to illustrate the complexity of the problem faced by any stereoscopic vision system (Marr 1982; Marr and Poggio 1976). Interestingly, complex cells, in the form of disparity-energy units, appear to provide a processing stage necessary for solving the problem. Not only are complex cells excited by stimuli that are matched correctly, point-by-point, for the two eyes, but they also are excited by stimulus combinations that are considered from the traditional point of view to be incorrect matches (bipartite circles in Fig. 22). And yet, this behavior is advantageous, as the situation inevitably occurs under most natural viewing con-

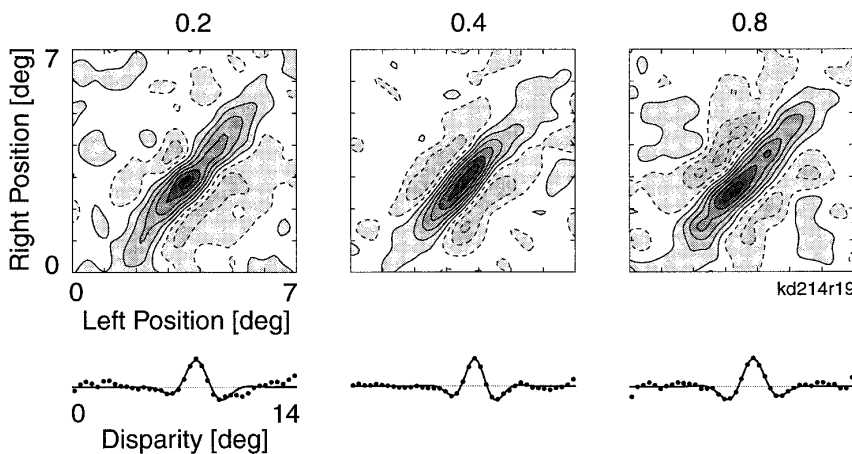


FIG. 21. Results of an experiment to test hypotheses of Fig. 20. Binocular RFs in X_L - X_R domain and disparity tuning curves are shown for 3 stimulus widths of 0.2, 0.4, and 0.8° . This cell was recorded from layer 2+3. Stimulus edge hypothesis fails.

ditions. In other words, false-matches (bipartite circles) should not necessarily be rejected but make an important contribution to identifying the correct overall match between local regions in the left and right images.

Given these considerations, the following picture emerges. Disparity sensors, as implemented by complex cells, solve the matching problem for a localized region of space. This is a partial solution that raises the level of primitives that are matched binocularly from individual white and black elements of images (as illustrated by Fig. 1) to small patches of image approximately the size of the RFs. Then, the matching problem still remains to be solved across these spatially distributed image patches. For a subsequent processing stage that receives the output of these complex cells, this reduces the complexity of the binocular matching problem (as measured by the number of false matches) by a factor of four to nine or more, given that at least two to three subregions typically are present in a subunit RF (DeAngelis et al. 1995a; Gaska et al. 1994; Movshon et al. 1978a; Szulborski and Palmer 1990). Therefore, the response of complex cells to opposite contrast stimuli does not compound the problem of ambiguous stereo matches (Blake and Wilson 1991). Rather, it contributes positively to the solution of the stereo matching problem. In retrospect, the original presentation of the correspondence problem (Julesz 1971; Marr and Poggio 1976) may have overemphasized the complexity of the problem, because it is likely that nowhere in the stereo processing stream, the system actually tries to match individual black and white elements in the two images.

Deviations of neural responses from predictions of the disparity energy model

Although the energy model provides a good description of the data from many complex cells (Figs. 11 and 12), various instances and degrees of deviation were found. For some complex cells, it appears that a single energy unit consisting of four simple subunits is not sufficient. For these cells, a diagonal region of excitation extends along the frontoparallel plane over a much longer distance than can be accounted for by a single energy unit (Figs. 11C and 13A). Additional energy units appear necessary to cover the large spatial extent of these neurons' RFs. The requirement for multiple energy units still places the overall scheme within the scope of the energy model. After all, a single energy unit model is the most parsimonious configuration that satisfies the properties that actual complex cells exhibit, and it is not surprising that neurons collect input from many more cells than the minimum configuration requires. In fact, such an extended spatial coverage is beneficial because it provides increased positional invariance while retaining narrow selectivity to disparity. There is some psychophysical evidence that humans may rely on such mechanisms (McKee et al. 1990). Recent computational studies also show that pooling activities of multiple energy units can eliminate false matches and reduce noise in disparity estimates (Fleet et al. 1996; Qian and Zhu 1997). Note that multiple energy units must be tuned to a single common disparity. This requires a remarkable degree of specificity of neural wiring, when one considers that the total number of subunits is four times the number of energy units. Also note that an extended spa-

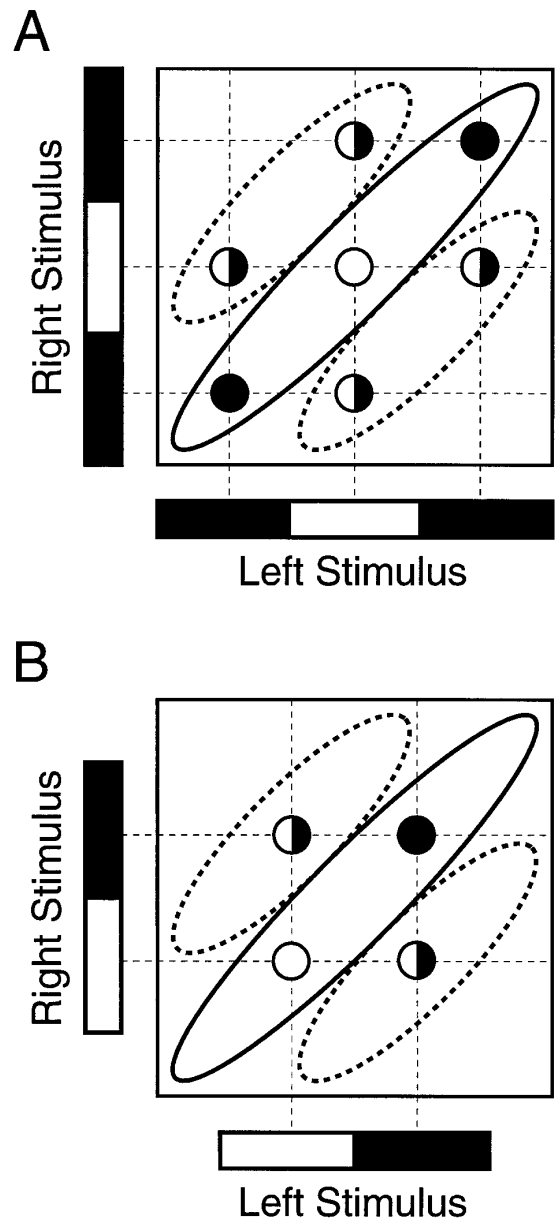


FIG. 22. Schematic illustrations are shown that describe benefits of responses to opposite contrast stimuli. *A*: a typical binocular RF in X_L - X_R domain is shown in relation to a likely stimulus configuration with multiple targets for each eye. Solid diagonal contour represents an excitatory subregion to matched contrast stimuli, while dashed contours depict excitatory regions to opposite contrast stimuli (see Fig. 15). \bullet , opposite contrast targets; \circ and \bullet , contrast matched targets. All targets fall within appropriate regions of binocular RF. Band-pass nature of linear subunits guarantees that stimulus periodicity, as shown, is predominant spatial frequency component for any natural complex scenes. *B*: same binocular RF is insensitive to exact monocular positions of stimuli as long as binocular disparity is appropriate for the cell.

tial region cannot be achieved by merely increasing the RF size of individual subunits in a single energy-unit model. If the subunit RF size is increased while the preferred spatial frequency is fixed, there will be additional spatial subregions within the RF. This should lead to a corresponding increase in the number of subregions (alternating excitatory and suppressive regions) in the disparity tuning curve. However, we do not observe such extra regions for the cell of Fig. 6B,

which clearly had a large spatial extent, as shown in Fig. 12C. Multiple energy units that are spatially distributed with partial overlap can provide a large spatial coverage without introducing additional ripples.

Another deviation from predictions of the energy model is the tendency for disparity frequency to be lower than optimal spatial frequency as measured monocularly with sinusoidal gratings (Fig. 17C). Assuming that the optimal spatial frequency represents the selectivity of linear subunits (Gaska et al. 1994; Movshon et al. 1978a; Szulborski and Palmer 1990), the energy model predicts a close match between the two. Interestingly, similar deviations have been found with monocular measurements between RF data and those from measurements with sinusoidal gratings. Optimal and cutoff spatial frequencies predicted from monocular second-order kernels were slightly lower (~ 0.25 octaves) than those measured with grating stimuli for complex cells (Gaska et al. 1994; Szulborski and Palmer 1991). Although the origin of the discrepancy between disparity frequency and spatial frequency is not clear, it is possible that the deviation that we find for disparity frequency shares a common basis with those found monocularly.

A linear regression of the data in Fig. 17C reveals a statistically significant correlation between the optimal spatial frequency and the disparity frequency. However, the slope is < 1 , which means that cells tuned to high spatial frequencies tend to have a lower disparity frequency than the energy model predicts. The physiological basis of this deviation is not clear. One possibility is that the deviation is due to gain normalization mechanisms that may operate at various stages of the energy model. A computational study indicates that gain normalization mechanisms are able to modify details of disparity tuning curves, e.g., by attenuating secondary peaks in disparity tuning curves (Fleet et al. 1995). Our current model is a strictly feed-forward version and does not include any gain normalization mechanisms. Based on prevalence of gain control phenomena observed in cortical neurons (Albrecht et al. 1984; Carandini and Heeger 1994; Heeger 1992a,b; Ohzawa et al. 1982, 1985), such mechanisms must clearly be incorporated.

Although we cannot speculate any further on possible causes of the disparity frequency shift, we note that there may be a psychophysical manifestation of this deviation. The deviation of the disparity frequency, to a lower value than is predictable by the optimal spatial frequency, becomes more pronounced as the spatial frequency increases (see the regression line in Fig. 17C). Because of an inverse relationship between the frequency and the period, the lower the disparity frequency, the larger the equivalent disparity range becomes for a given range of phase. This trend predicts a disparity range that is larger than that predicted from a strict phase model. Therefore this deviation is indeed consistent with the psychophysical finding that the binocular fusion range for band-pass filtered random-dot stereograms becomes larger than the range predicted by the phase model at high spatial frequencies (Smallman and MacLeod 1994). Other studies show a similar trend except that there is a relatively abrupt transition near the spatial frequency of 2.5 cycles/deg (Legge and Gu 1989; Schor and Wood 1983) instead of a gradual change as reported by Smallman and MacLeod (1994). We also should note that the disparity

range of complex cells, as defined by the extent of the disparity tuning curve (Fig. 17D), decreases with a slope substantially less than -1 as predicted by a strict phase model. This is also consistent with the psychophysical findings. It is likely that both of these factors, disparity frequency and disparity range, of neurons must be taken into account when attempting to relate psychophysics and physiology on the issue of size-disparity correlation.

Another type of deviation of cell responses from predictions of the energy model is an inconsistency regarding different combinations of the sign of stimulus contrast. We find cases in which the energy model provides an almost perfect fit for one combination of stimulus contrasts, while failing for others. Some of this type of deviation is probably attributable to a differential effectiveness of bright and dark stimuli for some complex cells. That is, it is not unusual to find cells that respond better to bright stimuli than to dark stimuli, or vice versa. However, there are deviations that cannot be explained by such asymmetries of responses to the sign of contrast. For example, the cell presented in Fig. 11C shows a reasonable fit for the dark-dark condition, though it probably needs more energy units. The cell also had nearly balanced responses to bright and dark stimuli (Fig. 6). However, the response to the bright-dark condition (Fig. 11D) exhibits hardly any binocular interaction. The energy model predicts that this condition should also cause an elongated excitatory region oriented at 45° but shifted to a different disparity. The binocular term for Fig. 11D should have been nearly as strong as that for Fig. 11C with phase inversion of the binocular term. The reason for this type of deviation is not clear.

Lack of sensitivity to motion-in-depth

RFs of most cells show no obvious signs of orientation (i.e., tilt) in the D-T domain (Fig. 15), indicating that the preferred disparity of the cell does not change over the time course of the response. A quantitative evaluation of responses in the frequency domain shows that cells are not sensitive to motion-in-depth (Cynader and Regan 1978, 1982; Spileers et al. 1990) as indicated by extremely low values of velocity-in-depth (Fig. 19A). These results may be related to a psychophysical finding that speed discrimination performance for targets moving in depth is very poor, and the task is possible only when the targets move slowly (Harris and Watamaniuk 1995). The lack of motion-in-depth sensitivity for complex cells is not surprising because none of the simple cells sampled in a previous study ($n = 65$) had opposite preferred directions of motion for the two eyes (DeAngelis et al. 1995a; Ohzawa et al. 1996). In addition, most cells maintained a constant preferred disparity over the time course of the response. Even for space-time inseparable cells, the rate of change in the RF phases was matched closely between the two eyes (DeAngelis et al. 1995a; Ohzawa et al. 1996). Preferred monocular velocities for the two eyes also are matched closely, indicating that direction-selective simple cells primarily encode information about motion within fronto-parallel planes (Ohzawa et al. 1996). If these simple cells serve as subunits for complex cells, a lack of tilt of RF orientation in the D-T domain is expected.

Results presented in this paper are based on first-order

binocular responses in the sense that we have measured responses to a single binocularly viewed target in three-dimensional space. Therefore, conceptually, our D-T plots (Figs. 15 and 20A) are analogous to the X-T plots obtained in monocular studies of simple cells (DeAngelis et al. 1993a, 1995a; McLean and Palmer 1989; McLean et al. 1994). In both cases, we measure RFs in response to a single target in real space. However, the procedural aspects of the analyses, presented herein for complex cells, involve computations of second-order responses, or interactions between two stimuli, because there are two stimuli for a binocular target, one for each eye. Although these X-T and D-T RFs provide rich information on cell responses to motion and motion-in-depth, respectively, they alone may not give us a complete picture. This is one limitation of the present study. A complete evaluation of monocular motion sensitivity requires motion stimuli consisting of at least two sequentially presented stimuli with spatial offsets (Emerson et al. 1987, 1992). Similarly, for binocular studies, it appears necessary to measure the interactions between two sequentially presented binocular targets in three-dimensional space. Although, conceptually, such stimuli are second-order in real space, nominally there will be a total of four stimuli for controlled dichoptic presentation. Thus, the analyses required will be fourth-order. Analyses of such high-order interactions may be extremely difficult even with modern nonlinear analysis techniques (Anzai et al. 1995; Sutter 1991).

How many disparity energy units are needed for a complete disparity representation?

What is the minimum number of cells required to implement a given neural computation? This is an important question for a number of reasons. Although there appears to be an abundance of neurons in the visual cortex, it seems reasonable to assume that the brain encodes information in an efficient manner. Many aspects of visual information encoding seem to be designed for high efficiency, such as the representation of binocular information by a population of simple cells (DeAngelis et al. 1991, 1995a; Ohzawa et al. 1996), and the representation of space-time information in the LGN (Dan et al. 1996). We have shown for the disparity-energy model that a minimum of four simple subunits is necessary to build a disparity-sensitive complex cell (Fig. 8A). If these subunits are capable of signalling negative values (i.e., inhibition), then two linear subunits will suffice (Fleet et al. 1996, 1997). In general, image representation schemes based on decompositions into wavelets and Gabor-like orthogonal basis functions require two neurons that can transmit bipolar signals (positive and negative) for each location, spatial frequency, and orientation (Daugmann 1985; Geisler and Hamilton 1986; Robson 1983; Sakitt and Barlow 1982; Watson 1983; Watson and Ahumada 1989, 1991). Because simple cells cannot signal negative values in their spike discharges due to a lack of spontaneous activity, a push-pull configuration requires double the number of cells (Pollen and Ronner 1981; Pollen et al. 1989). An array of these simple cells is thought to form a coarse-to-fine binocular image representation for encoding disparity information as well as form information (Marr and Poggio 1979; Ohzawa et al. 1996; Wilson et al. 1991).

Applying analogous considerations to the representation of depth, we ask how many complex cells (energy units) are required per spatial frequency, location, and orientation for a complete representation of binocular disparity. According to the general encoding rule mentioned above, this number should be two. *Equations 5 and 6* show that an energy unit carries a strictly positive signal (a sum of squares is always positive). Therefore there is no need to double the number to construct a push-pull organization. The binocular term of *Eq. 7* carries a bipolar signal, but the two monocular terms provide a conditioned excitation level about which responses may be modulated by the binocular term. A recent computational study also shows that only two complex cells per position are needed for reliable estimation of disparity (Qian and Zhu 1997). Note that the complex cell stage, as modeled, has no inhibitory input that subtracts linearly from the converging excitatory input from the subunits. Therefore the suppression that we mentioned, when describing the binocular response that is driven below the monocular excitation levels (Figs. 3A, 5, and 6), is an indirect effect of reduced responses of the simple subunits, which then are carried over to the complex cell stage. Thus we have avoided the term "inhibition" in referring to the reduction of a complex cell's response to levels below the monocular excitation. However, this does not mean that inhibition is not involved in constructing these complex cell RFs. Intracellular recordings from complex cells clearly show strong inhibitory postsynaptic potentials to visual stimuli as well as to electrical stimulation in the LGN (Ferster 1986). As we note above, gain normalization and contrast gain control mechanisms must rely on some form of inhibition, possibly both divisive and subtractive ones (Carandini and Ferster 1996; Carandini et al. 1996; Heeger 1992a,b). Inhibition that is involved in these mechanisms is likely to operate at the complex cell stage as well as for simple cells (Fleet et al. 1995, 1996b).

Disparity-insensitive complex cells

About 40% of complex cells are not tuned for binocular disparity (Hammond 1991; Ohzawa and Freeman 1986b). Data from only one such neuron are presented in this paper (Fig. 7), because our primary focus was on disparity-sensitive cells. The role of these cells remains a matter of speculation. Their output may provide a signal that is useful for motion and texture detection as well as contrast gain control and gain normalization (Heeger 1992a). These are functions that do not require binocular input.

On the other hand, it is possible that these nondisparity-selective cells play a role in stereopsis. Recalling *Eq. 7* and Fig. 10, we note that nondisparity-sensitive cells provide a signal that is the sum of monocular responses, as represented by the first two terms of *Eq. 7*. By subtracting the output of a nondisparity-sensitive cell from that of a disparity-sensitive cell with the same RF position, it is possible to compute the pure binocular interaction component represented by the third term of *Eq. 7*. This term represents a binocular cross-correlation operation that may be important for stereopsis (Fleet et al. 1996). Regardless of whether there is a neuron that actually computes the difference, it is clear that signals for computing binocular correlation are readily available within the striate cortex.

Hierarchical models and roles of simple and complex cells

The energy model that we have proposed has a hierarchical structure in which signals from a specific set of simple cell like subunits feed into a complex cell. Many complex cells respond almost exactly as predicted by the energy model, although some deviations are found. Although, the hierarchical organization of disparity-sensitive complex cells is quite likely, the question of whether these subunits are actually simple cells still remains open. For example, it is possible to attribute the subunits to a part of the complex cell structure, such as a portion of a dendrite that may operate as an independent integration unit (Mel 1993; Shepherd 1996). This possibility is reinforced by two factors. First, the distribution of the phase of the disparity tuning curves for our sample of complex cells (Fig. 17A) does not follow the asymmetry in the distribution for simple cells (DeAngelis et al. 1991; Ohzawa et al. 1996). If the complex cell RF is constructed from the simple cells' output according to the energy model, the two distributions should be closely similar. Second, the disparity frequency tended to be lower than the monocular preferred spatial frequency (Fig. 17C). Again, if the hierarchical energy model is correct, the two frequencies should be the same. There are two possibilities for the cause of these discrepancies: one is that the hierarchical assumption that signals flow from simple to complex cells is not correct. The other is that the energy model is incorrect in that complex cells may combine simple cell output using an entirely different scheme. Unfortunately, it is not possible to determine, based on our current knowledge, which of these factors contribute to the discrepancies. Therefore, although it appears unlikely that the visual system is built in an inefficient manner by which complex cells duplicate identical computations that are performed by simple cells, there is a clear lack of direct evidence for monosynaptic connections from simple cells to complex cells (Ghose et al. 1994; Toyama et al. 1981; but see Alonso 1996; Liu 1993). Direct LGN input to complex cells also has been reported (Bullier and Henry 1979a–c; Henry et al. 1979; Hoffmann and Stone 1971; Stone 1972; Tanaka 1983). However, such input does not necessarily contradict a basic hierarchical structure, i.e., both hierarchical and direct LGN input may be present simultaneously. Nevertheless, we must conclude that the precise nature of the RF subunits of complex cells is still unknown.

Although the exact details of complex cell circuitry are yet to be worked out, we now know a great deal about how to build a unit that is functionally equivalent to a real neuron. In other words, we are at a point where we can provide an equivalent circuit diagram for complex cells, if not an actual one. The motion energy model (Adelson and Bergen 1985; Emerson et al. 1992; Pollen et al. 1989) provides a good functional schematic of complex cells for monocular motion and motion parallel to frontoparallel planes. Similarly, the disparity energy model described here presents a reasonable functional circuitry of complex cells for stereopsis. The fact that a computational model based on an array of disparity energy units can solve dynamic random dot stereograms (Qian 1994) provides strong support for the scheme that we have proposed.

In our previous paper on encoding of binocular information by simple cells (Ohzawa et al. 1996), we emphasized that simple cells should not be considered to be playing an exclusive role for a specific visual function, such as stereopsis or form perception, because the information they carry may be used for a variety of other perceptual tasks. Instead of trying to associate simple cells with a specific function such as stereopsis, we have described the notion that, as a population, simple cells encode nearly complete but uncommitted information, via a binocular linear transform, which may be used for any purpose, including, but not limited to, stereo, motion, and form perception. Simple cells may be selected according to appropriate sets of constraints to produce second stage neurons in the visual cortex. Complex cells appear to be the next stage of processing for specific binocular tasks, because disparity-sensitive complex cells must be collecting input from a set of simple cells that share a common preferred disparity. Given the possibility that many simple cells may feed into a complex cell, this is a remarkably tight constraint, which apparently is satisfied for many cells. Similarly, complex cells also must have a specific organization of their subunits to function as motion energy sensors (Adelson and Bergen 1985; Emerson et al. 1992). The rules for selecting appropriate simple cells for a motion sensor are similar to those for disparity energy sensors, and there is also a set of strict constraints for the selection. As with simple cells, however, it probably would be a mistake to divide the complex cell population into a group that is responsible for stereopsis only and another that is responsible for motion processing only. Most likely, the same set of complex cells performs computations for both functions at the same time. A computational modeling study shows that such an integrated model of motion-stereo representation is indeed possible (Qian 1994; Qian and Andersen 1996).

In conclusion, we have presented results of detailed measurements of binocular responses from complex cells, and comparisons of the data with predictions of the disparity energy model. There is generally good agreement between the data and the model predictions. Combined with the results from simple cells (Ohzawa et al. 1996) and other studies, we now have a reasonable functional schematic diagram of the early stages of binocular visual information processing, as well as a clearer picture of the roles that simple and complex cells play in the striate cortex.

We thank G. Ghose for help with the experiments and A. Anzai for discussions and for suggesting the use of composite profiles in Fig. 15. We also thank Drs. D. J. Fleet and H. S. Smallman and two anonymous reviewers for valuable comments on the manuscript.

This work was supported by National Eye Institute research and CORE grants (EY-01175 and EY-03176) and by the Human Frontier Science Program. Additional materials related to this study are available via the World Wide Web at <http://totoro.berkeley.edu/izumi/stereopsis/>.

Present address of G. C. DeAngelis: Dept. of Neurobiology, Stanford University School of Medicine, Stanford, CA 94305-5401.

Address for reprint requests: I. Ohzawa, School of Optometry, University of California, 360 Minor Hall, Berkeley, CA 94720-2020.

Received 7 November 1996; accepted in final form 4 February 1997.

REFERENCES

- ADELSON, E. H. AND BERGEN, J. R. Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am.* 2: 284–299, 1985.

- ALBRECHT, D. G., FARRAR, S. B., AND HAMILTON, D. B. Spatial contrast adaptation characteristics of neurones recorded in the cat's visual cortex. *J. Physiol. Lond.* 347: 713–739, 1984.
- ALONSO, J. M. The microcircuitry of complex cells in cat striate cortex. *Soc. Neurosci. Abstr.* 22: 489, 1996.
- ANZAI, A., OHZAWA, I., FREEMAN, R. D., AND COHN, T. E. Do simple cells in the cat's striate cortex encode binocular disparity through position and phase incongruities? *Soc. Neurosci. Abstr.* 21: 1648, 1995.
- BAKER, C. L., JR. Spatial- and temporal-frequency selectivity as a basis for velocity preference in cat striate cortex neurons. *Visual Neurosci.* 4: 101–113, 1990.
- BARLOW, H. B., BLAKEMORE, C., AND PETTIGREW, J. D. The neural mechanism of binocular depth discrimination. *J. Physiol. Lond.* 193: 327–342, 1967.
- BLAKE, R. AND WILSON, H. R. Neural models of stereoscopic vision. *Trends Neurosci.* 14: 445–452, 1991.
- BRACEWELL, R. N. *The Fourier Transform and its Applications* (2nd ed.). New York: McGraw-Hill, 1978.
- BULLIER, J. AND HENRY, G. H. Ordinal position of neurons in cat striate cortex. *J. Neurophysiol.* 42: 1251–1263, 1979a.
- BULLIER, J. AND HENRY, G. H. Neural path taken by afferent streams in striate cortex of the cat. *J. Neurophysiol.* 42: 1264–1270, 1979b.
- BULLIER, J. AND HENRY, G. H. Laminar distribution of first-order neurons and afferent terminals in cat striate cortex. *J. Neurophysiol.* 42: 1271–1281, 1979c.
- BURR, D. AND ROSS, J. Visual processing of motion. *Trends Neurosci.* 9: 304–307, 1986.
- CARANDINI, M. AND FERSTER, D. Visual adaptation hyperpolarizes cells of CARANDINI, M. AND HEEGER, D. J. Summation and division by neurons in primate visual cortex. *Science Wash. DC* 264: 1333–1336, 1994.
- CARANDINI, M., HEEGER, D. J., AND MOVSHON, J. A. Linearity and gain control in V1 simple cells. In: *Cerebral Cortex, Cortical Models*, edited by E. G. Jones and P. S. Ulinski. New York: Plenum, In press.
- CHINO, Y. M., SMITH, E. L., YOSHIDA, K., CHENG, H., AND HAMAMOTO, J. Binocular interactions in striate cortical neurons of cats reared with discordant visual inputs. *J. Neurosci.* 14: 5050–5067, 1994.
- CYNADER, M. AND REGAN, D. Neurones in cat parastriate cortex sensitive to the direction of motion in three-dimensional space. *J. Physiol. Lond.* 274: 549–569, 1978.
- CYNADER, M. AND REGAN, D. Neurones in cat visual cortex tuned to the direction of motion in depth: effect of positional disparity. *Vision Res.* 22: 967–982, 1982.
- DAN, Y., ATICK, J. J., AND REID, R. C. Efficient coding of natural scenes in the lateral geniculate nucleus—experimental test of a computational theory. *J. Neurosci.* 16: 3351–3362, 1996.
- DAUGMANN, J. G. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *J. Opt. Soc. Am.* 2: 1160–1169, 1985.
- DEANGELIS, G. C., OHZAWA, I., AND FREEMAN, R. D. Depth is encoded in the visual cortex by a specialized receptive field structure. *Nature Lond.* 352: 156–159, 1991.
- DEANGELIS, G. C., OHZAWA, I., AND 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, 1993a.
- DEANGELIS, G. C., OHZAWA, I., AND FREEMAN, R. D. 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.
- DEANGELIS, G. C., OHZAWA, I., AND FREEMAN, R. D. Neural mechanisms underlying stereopsis: how do simple cells in the visual cortex encode binocular disparity? *Perception* 24: 3–31, 1995a.
- DEANGELIS, G. C., OHZAWA, I., AND FREEMAN, R. D. Receptive-field dynamics in the central visual pathways. *Trends Neurosci.* 18: 451–458, 1995b.
- DEBOER, E. AND KUYPER, P. Triggered correlation. *IEEE Trans. Biomed. Eng.* 15: 169–179, 1968.
- DEVALOIS, R. L., ALBRECHT, D. G., AND THORELL, L. G. Spatial frequency selectivity of cells in macaque visual cortex. *Vision Res.* 22: 545–559, 1982.
- DOUGLAS, R. J. AND MARTIN, K. A. C. A functional microcircuit for cat visual cortex. *J. Physiol. Lond.* 440: 735–769, 1991.
- EGGERMONT, J. J., JOHANNESMA, P. I. M., AND AERTSEN, A. M. H. J. Reverse-correlation methods in auditory research. *Q. Rev. Biophys.* 16: 341–414, 1983.
- EMERSON, R. C., CITRON, M. C., VAUGHN, W. J., AND KLEIN, S. A. Nonlinear directionally selective subunits in complex cells of cat striate cortex. *J. Neurophysiol.* 58: 33–65, 1987.
- EMERSON, R. C., KORENBERG, M. J., AND CITRON, M. C. Identification of intensive nonlinearities in cascade models of visual cortex and its relation to cell classification. In: *Advanced Methods of Physiological System Modelling*, edited by V. Z. Marmarelis. New York: Plenum, 1989, vol. 2, p. 97–111.
- EMERSON, R. C., BERGEN, J. R., AND ADELSON, E. H. Directionally selective complex cells and the computation of motion energy in cat visual cortex. *Vision Res.* 32: 203–218, 1992.
- FERSTER, D. A comparison of binocular depth mechanisms in Areas 17 and 18 of the cat visual cortex. *J. Physiol. Lond.* 311: 623–655, 1981.
- FERSTER, D. Orientation selectivity of synaptic potentials in neurons of cat primary visual cortex. *J. Neurosci.* 6: 1284–1301, 1986.
- FIELD, D. J. AND TOLHURST, D. J. The structure and symmetry of simple-cell receptive-field profiles in the cat's visual cortex. *Proc. R. Soc. Lond. B Biol. Sci.* 228: 379–400, 1986.
- FLEET, D. J., HEEGER, D. J., AND WAGNER, H. Computational model of binocular disparity (Abstract). *Invest. Ophthalmol. Visual Sci.* 36, Suppl.: 365, 1995.
- FLEET, D. J., HEEGER, D. J., AND WAGNER, H. Modelling binocular neurons in the primary visual cortex. In: *Computational and Biological Mechanisms of Visual Coding*, edited by M. Jenkin and L. Harris. Cambridge, UK: Cambridge Univ. Press, In press.
- FLEET, D. J., WAGNER, H., AND HEEGER, D. J. Encoding of binocular disparity: energy models, position shifts and phase shifts. *Vision Res.* 36: 1839–1857, 1996.
- FREEMAN, R. D. AND OHZAWA, I. On the neurophysiological organization of binocular vision. *Vision Res.* 30: 1661–1676, 1990.
- GABOR, D. Theory of communication. *J. Inst. Elec. Eng.* 93: 429–457, 1946.
- GASKA, J. P., JACOBSON, L. D., CHEN, H. W., AND POLLEN, D. A. Space-time spectra of complex cell filters in the macaque monkey: a comparison of results obtained with pseudowhite noise and grating stimuli. *Visual Neurosci.* 11: 805–821, 1994.
- GASKA, J. P., POLLEN, D. A., AND CAVANAGH, P. Diversity of complex cell responses to even- and odd-symmetric luminance profiles in the visual cortex of the cat. *Exp. Brain Res.* 68: 249–259, 1987.
- GEISLER, W. S. AND HAMILTON, D. B. Sampling-theory analysis of spatial vision. *J. Opt. Soc. Am.* 3: 62–70, 1986.
- GHOSE, G. M., FREEMAN, R. D., AND OHZAWA, I. Local intracortical connections in the cat's visual cortex: postnatal development and plasticity. *J. Neurophysiol.* 72: 1290–1303, 1994.
- GHOSE, G. M., OHZAWA, I., AND FREEMAN, R. D. A flexible PC-based physiological monitor for animal experiments. *J. Neurosci. Methods* 62: 7–13, 1995.
- GIZZI, M. S., KATZ, E., SCHUMER, R. A., AND MOVSHON, J. A. Selectivity for orientation and direction of motion of single neurons in cat striate and extrastriate visual cortex. *J. Neurophysiol.* 63: 1529–1543, 1990.
- HAMMOND, P. Binocular phase specificity of striate cortical neurones. *Exp. Brain Res.* 87: 615–623, 1991.
- HARRIS, J. M. AND WATAMANIUK, S. N. Speed discrimination of motion-in-depth using binocular cues. *Vision Res.* 35: 885–896, 1995.
- HEEGER, D. Normalization of cell responses in cat striate cortex. *Visual Neurosci.* 9: 181–198, 1992a.
- HEEGER, D. Half-squaring in responses of cat striate cells. *Visual Neurosci.* 9: 427–443, 1992b.
- HENRY, G. H., HARVEY, A. R., AND LUND, J. S. The afferent connections and laminar distribution of cells in the cat striate cortex. *J. Comp. Neurol.* 187: 725–744, 1979.
- HOFFMANN, K. P. AND STONE, J. Conduction velocity of afferents to cat visual cortex: a correlation with cortical receptive field properties. *Brain Res.* 32: 460–466, 1971.
- HOWARD I. P. AND ROGERS, B. J. *Binocular Vision and Stereopsis*. New York: Oxford Univ. Press, 1995.
- HUBEL, D. H. AND WIESEL, T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol. Lond.* 160: 106–154, 1962.
- HUBEL, D. H. AND WIESEL, T. N. Receptive fields and functional architecture of monkey striate cortex. *J. Physiol. Lond.* 195: 215–243, 1968.
- JACOBSON, L. D., GASKA, J. P., AND POLLEN, D. A. Phase, displacement,

- and hybrid models for disparity coding (Abstract). *Invest. Ophthalmol. Visual Sci.* 34, Suppl.: 908, 1993.
- JONES, J. P. AND PALMER, L. A. The two-dimensional spatial structure of simple receptive fields in the cat striate cortex. *J. Neurophysiol.* 58: 1187–1211, 1987a.
- JONES, J. P. AND PALMER, L. A. An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *J. Neurophysiol.* 58: 1233–1258, 1987b.
- JOSHUA, D. E. AND BISHOP, P. O. Binocular single vision and depth discrimination. Receptive field disparities for central and peripheral vision and binocular interaction on peripheral single units in cat striate cortex. *Exp. Brain Res.* 10: 389–416, 1970.
- JULESZ, B. Binocular depth perception of computer-generated patterns. *Bell Syst. Tech. J.* 39: 1125–1162, 1960.
- JULESZ, B. Experiment in perception. *Psychol. Today* 2: 16–23, 1968.
- JULESZ, B. *Foundations of Cyclopean Perception*. Chicago, IL: Univ. of Chicago Press, 1971.
- LEGGE, G. E. AND GU, Y. Stereopsis and contrast. *Vision Res.* 29: 989–1004, 1989.
- LEVAY, S. AND VOIGT, T. Ocular dominance and disparity coding in cat visual cortex. *Visual Neurosci.* 1: 395–414, 1988.
- LEVICK, W. R. Another tungsten microelectrode. *Med. Biol. Eng.* 10: 510–515, 1972.
- LIU, Z. *Functional microcircuitry of the striate cortex: a cross-correlation study* (PhD thesis). Cambridge, MA: Harvard Univ., 1993.
- MARCELJÀ, S. Mathematical description of the responses of simple cortical cells. *J. Opt. Soc. Am.* 70: 1297–1300, 1980.
- MARR, D. *Vision*. New York: W. H. Freeman, 1982.
- MARR, D. AND POGGIO, T. Cooperative computation of stereo disparity. *Science Wash. DC* 194: 283–287, 1976.
- MARR, D. AND POGGIO, T. A computational theory of human stereo vision. *Proc. R. Soc. Lond. B Biol. Sci.* 204: 301–328, 1979.
- MASKE, R., YAMANE, S., AND BISHOP, P. O. Binocular simple cells for local stereopsis: comparison of receptive field organizations for the two eyes. *Vision Res.* 24: 1921–1929, 1984.
- McKEE, S. P., LEVI, D. M., AND BOWNE, S. F. The imprecision of stereopsis. *Vision Res.* 30: 1763–1779, 1990.
- McLEAN, J. AND PALMER, L. A. Contribution of linear spatiotemporal receptive field structure to velocity selectivity of simple cells in the cat's striate cortex. *Vision Res.* 29: 675–679, 1989.
- McLEAN, J., RAAB, S., AND PALMER, L. A. Contribution of linear mechanisms to the specification of local motion by simple cells in areas 17 and 18 of the cat. *Visual Neurosci.* 11: 271–294, 1994.
- MEL, B. W. Synaptic integration in an excitable dendritic tree. *J. Neurophysiol.* 70: 1086–1101, 1993.
- MOVSHON, J. A. The velocity tuning of single units in cat striate cortex. *J. Physiol. Lond.* 249: 445–468, 1975.
- MOVSHON, J. A., THOMPSON, I. D., AND TOLHURST, D. J. Receptive field organization of complex cells in the cat's striate cortex. *J. Physiol. Lond.* 283: 79–99, 1978a.
- MOVSHON, J. A., THOMPSON, I. D., AND TOLHURST, D. J. Spatial and temporal contrast sensitivity of neurones in areas 17 and 18 of the cat's visual cortex. *J. Physiol. Lond.* 283: 101–120, 1978b.
- NIKARA, T., BISHOP, P. O., AND PETTIGREW, J. D. Analysis of retinal correspondence by studying receptive fields of binocular single units in cat striate cortex. *Exp. Brain Res.* 6: 353–372, 1968.
- OHZAWA, I., DEANGELIS, G. C., AND FREEMAN, R. D. Stereoscopic depth discrimination in the visual cortex: neurons ideally suited as disparity detectors. *Science Wash. DC* 249: 1037–1041, 1990.
- OHZAWA, I., DEANGELIS, G. C., AND FREEMAN, R. D. Overlap of subunits of complex cell receptive fields. *Soc. Neurosci. Abstr.* 21: 22, 1995.
- OHZAWA, I., DEANGELIS, G. C., AND FREEMAN, R. D. Encoding of binocular disparity by simple cells in the cat's visual cortex. *J. Neurophysiol.* 75: 1179–1805, 1996.
- OHZAWA, I., AND FREEMAN, R. D. The binocular organization of simple cells in the cat's visual cortex. *J. Neurophysiol.* 56: 221–242, 1986a.
- OHZAWA, I., AND FREEMAN, R. D. The binocular organization of complex cells in the cat's visual cortex. *J. Neurophysiol.* 56: 243–259, 1986b.
- OHZAWA, I., SCLAR, G., AND FREEMAN, R. D. Contrast gain control in the cat visual cortex. *Nature Lond.* 298: 266–268, 1982.
- OHZAWA, I., SCLAR, G., AND FREEMAN, R. D. Contrast gain control in the cat's visual system. *J. Neurophysiol.* 54: 651–667, 1985.
- PETTIGREW, J. D., NIKARA, T., AND BISHOP, P. O. Binocular interaction on single units in cat striate cortex: simultaneous stimulation by single moving slit with receptive fields in correspondence. *Exp. Brain Res.* 6: 391–410, 1968.
- PIERCE, D. M. AND BENTON, A. L. Relationship between monocular and binocular depth acuity. *Ophthalmologica* 170: 43–50, 1975.
- POGGIO, G. F. Mechanisms of stereopsis in monkey visual cortex. *Cereb. Cortex* 3: 193–204, 1995.
- POGGIO, G. F. AND FISCHER, B. Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving Rhesus monkey. *J. Neurophysiol.* 40: 1392–1405, 1977.
- POGGIO, G. F., GONZALEZ, F., AND KRAUSE, F. Stereoscopic mechanisms in monkey visual cortex: binocular correlation and disparity selectivity. *J. Neurosci.* 8: 4531–4550, 1988.
- POGGIO, G. F., MOTTER, B. C., SQUATRITO, S., AND TROTTER, Y. Responses of neurons in visual cortex (V1 and V2) of the alert macaque to dynamic random-dot stereograms. *Vision Res.* 25: 397–406, 1985.
- POGGIO, G. F. AND POGGIO, T. The analysis of stereopsis. *Annu. Rev. Neurosci.* 7: 379–412, 1984.
- POLLEN, D. A., GASKA, J. P., AND JACOBSON, L. D. Physiological constraints on models of visual cortical function. In: *Models of Brain Function*, edited by R. M. J. Cotterill. Cambridge, UK: Cambridge Univ. Press, 1989, p. 115–135.
- POLLEN, D. A. AND RONNER, S. F. Phase relationships between adjacent simple cells in the visual cortex. *Science Wash. DC* 212: 1409–1411, 1981.
- PRESS, W. H., FLANNERY, B. P., TEUKOLSKY, S. A., AND VETTERLING, W. T. *Numerical Recipes in C* (2nd ed.). Cambridge, UK: Cambridge Univ. Press, 1992.
- QIAN, N. Computing stereo disparity and motion with known binocular cell properties. *Neural Comput.* 6: 390–404, 1994.
- QIAN, N. AND ANDERSEN, R. A. A physiological model for motion-stereo integration and a unified explanation of the Pulfrich-like phenomena. *Vision Res.* In press.
- QIAN, N. AND ZHU, Y. Physiological computation of binocular disparity. *Vision Res.* In press.
- REID, R. C., SOODAK, R. E., AND SHAPLEY, R. M. Directional selectivity and spatiotemporal structure of receptive fields of simple cells in cat striate cortex. *J. Neurophysiol.* 66: 505–529, 1991.
- ROBSON, J. G. Frequency domain visual processing. In: *Physical and Biological Processing of Images*, edited by O. J. Braddick and A. C. Sleight. Berlin: Springer-Verlag, 1983.
- SAKITT, B. AND BARLOW, H. B. A model for the economical encoding of the visual image in cerebral cortex. *Biol. Cybern.* 43: 97–108, 1982.
- SCHOR, C. M. AND WOOD, I. Disparity range for local stereopsis as a function of luminance spatial frequency. *Vision Res.* 23: 1649–1654, 1983.
- SHEPHERD, G. M. The dendritic spine: a multifunctional integrative unit. *J. Neurophysiol.* 75: 2197–2210, 1996.
- SKOTTUN, B. C., DEVALOIS, R. L., GROSOFF, D. H., MOVSHON, J. A., ALBRECHT, D. G., AND BONDS, A. B. Classifying simple and complex cells on the basis of response modulation. *Vision Res.* 31: 1079–1086, 1991.
- SKOTTUN, B. C., AND FREEMAN, R. D. Stimulus specificity of binocular cells in the cat's visual cortex: ocular dominance and the matching of left and right eyes. *Exp. Brain Res.* 56: 206–216, 1984.
- SMALLMAN, H. S. AND MACLEOD, D. I. A. A size-disparity correlation in stereopsis at contrast threshold. *J. Opt. Soc. Am.* 11: 2169–2183, 1994.
- SPILEERS, W., ORBAN, G. A., GULYAS, B., AND MAES, H. Selectivity of cat Area 18 neurons for direction and speed in depth. *J. Neurophysiol.* 63: 936–954, 1990.
- SPITZER, H. AND HOCHSTEIN, S. A complex cell receptive field model. *J. Physiol. Lond.* 53: 1266–1285, 1985.
- STONE, J. Morphology and physiology of the geniculo-cortical synapse in the cat: the question of parallel input to the striate cortex. *Invest. Ophthalmol. Visual Sci.* 11: 338–346, 1972.
- SUTTER, E. E. Electrophysiological investigation of visual units in the cat: a revised conception of receptive fields. *Soc. Neurosci. Abstr.* 4: 444, 1974.
- SUTTER, E. E. A revised conception of visual receptive fields based on pseudorandom spatio-temporal pattern stimuli. *Proc. Symp. Test. Ident. Nonlinear Syst.* 1: 353–365, 1975.
- SUTTER, E. E. The fast m-transform: a fast computation of cross-correlations with binary m-sequences. *SIAM J. Comput.* 20: 686–694, 1991.
- SZULBORSKI, R. G. AND PALMER, L. A. The two-dimensional spatial structure of nonlinear subunits in the receptive fields of complex cells. *Vision Res.* 30: 249–254, 1990.
- SZULBORSKI, R. G. AND PALMER, L. A. Linear behavior of complex cell subunits in cat striate cortex (Abstract). *Invest. Ophthalmol. Vis. Sci.* 32, Suppl.: 1253, 1991.

- TANAKA, K. Cross-correlation analysis of geniculostriate neuronal relationships in cats. *J. Neurophysiol.* 49: 1303–1318, 1983.
- TOYAMA, K., KIMURA, M., AND TANAKA, K. Organization of cat visual cortex as investigated by cross-correlation technique. *J. Neurophysiol.* 46: 202–214, 1981.
- VON DER HEYDT, R., ADORJANI, C. S., HÄNNY, P., AND BAUMGARTNER, G. Disparity sensitivity and receptive field incongruity of units in the cat striate cortex. *Exp. Brain Res.* 31: 523–545, 1978.
- WATSON, A. B. Detection and recognition of simple spatial forms. In: *Physical and Biological Processing of Images*, edited by O. J. Braddick and A. C. Slade. Berlin: Springer-Verlag, 1983.
- WATSON, A. B. Cortical organization. In: *Vision: Coding and Efficiency*, edited by C. B. Blakemore. Cambridge, UK: Cambridge Univ. Press, 1991.
- WATSON, A. B. AND AHUMADA, A. J. JR. Model of human visual motion sensing. *J. Opt. Soc. Am.* 2: 322–342, 1985.
- WATSON, A. B. AND AHUMADA, A. J., JR. A hexagonal orthogonal-oriented pyramid as a model of image representation in visual cortex. *IEEE Trans. Biomed. Eng.* 36: 97–106, 1989.
- WHEATSTONE, C. Contributions to the physiology of vision. I. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philos. Trans. R. Soc. Lond.* 2: 371–393, 1838.
- WILSON, H. R., BLAKE, R., AND HALPERN, D. L. Coarse spatial scales constrain the range of binocular fusion of fine scales. *J. Opt. Soc. Am.* 8: 229–236, 1991.
- ZHU, Y. D. AND QIAN, N. Binocular receptive field models, disparity tuning, and characteristic disparity. *Neural Comput.* 8: 1611–1641, 1996.