Spatial Frequency Integration for Binocular Correspondence in Macaque Area V4

Hironori Kumano, Seiji Tanabe, and Ichiro Fujita

Laboratory for Cognitive Neuroscience, Graduate School of Frontier Bioscience, and Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka, Japan

Kumano H, Tanabe S, Fujita I. Spatial frequency integration for binocular correspondence in macaque area V4. *J Neurophysiol* 99: 402–408, 2008. First published October 24, 2007; doi:10.1152/jn.00096.2007. Neurons in the primary visual cortex (V1) detect binocular disparity by computing the local disparity energy of stereo images. The representation of binocular disparity in V1 contradicts the global correspondence when the image is binocularly anticorrelated. To solve the stereo correspondence problem, this rudimentary representation of stereoscopic depth needs to be further processed in the extrastriate cortex. Integrating signals over multiple spatial frequency channels is one possible mechanism supported by theoretical and psychophysical studies. We examined selectivities of single V4 neurons for both binocular disparity and spatial frequency in two awake, fixating monkeys. Disparity tuning was examined with a binocularly correlated random-dot stereogram (RDS) as well as its anticorrelated counterpart, whereas spatial frequency tuning was examined with a sine wave grating or a narrowband noise. Neurons with broader spatial frequency tuning exhibited more attenuated disparity tuning for the anticorrelated RDS. Additional rectification at the output of the energy model does not likely account for this attenuation because the degree of attenuation does not differ among the various types of disparity-tuned neurons. The results suggest that disparity energy signals are integrated across spatial frequency channels for generating a representation of stereoscopic depth in V4.

INTRODUCTION

Identifying which features in the right retinal image corresponds to which features in the left retinal image, the stereo correspondence problem, is one of the most challenging tasks confronting the visual system (Julesz 1971; Marr and Poggio 1979). The problem is particularly prominent when viewing a random-dot stereogram (RDS) because each of the numerous dots dispersed in one retinal image can be matched with any dot in the other (Fig. 1*A*). A stereoscopic system can be tested with binocularly anticorrelated RDSs (i.e., contrast-reversed between the left and right images) to determine whether the system represents the global stereo correspondence (Fig. 1*A*). If the system represents the global correspondence, it should be insensitive to disparity in anticorrelated RDSs because there is no globally consistent match between the right and left eyes' dot patterns in these RDSs.

Although the stereo correspondence problem was originally formalized as searching for matching patterns, biological stereoscopic systems deal with the problem differently. Neurons at the initial, primary cortical (V1) stage compute the disparity energy of binocular images (Ohzawa et al. 1990). The stereo information encoded in the disparity-energy signal is not necessarily consistent with stereoscopic depth. For example, when the RDS is anticorrelated, disparity is encoded in a sign-inverted signal (Cumming and Parker 1997; see also Ohzawa et al. 1990), whereas stereo percept is abolished (Cogan et al. 1993; Cumming et al. 1998; Read and Eagle 2000). The abolished percept is accounted for by neural representations in higher cortical areas, where the encoding of disparity is lost for anticorrelated RDSs (Tanabe et al. 2004 for area V4; Janssen et al. 2003 for the inferior temporal cortex). The computation of stereoscopic information advances along the ventral visual pathway by attenuating neural responses initially elicited in V1 in the absence of a global match. How this computation is accomplished is unclear.

The disparity energy model states that V1 neurons encode the interocular phase difference of the spatial frequency (SF) components of the retinal image passing through the left and right eye's band-pass receptive fields (Ohzawa et al. 1990). Integration of neuronal signals across multiple SF bands helps to solve the correspondence problem (Fleet et al. 1996). A similar computation quantitatively explains the psychophysical performance of human stereo discrimination (Read 2002; Read and Eagle 2000). For binocular fusion, SF channels are not summed linearly, but integrated nonlinearly, as revealed by interaction between different SF bands (Rohaly and Wilson 1993; Smallman 1995; Wilson et al. 1991).

These studies promote neurophysiological investigations to test whether cortical neurons integrate over SF channels for stereo computation. V1 neurons integrate the outputs of several initial units by unevenly weighting different pass-bands (Menz and Freeman 2004a,b). However, whether this SF integration contributes to the attenuation of the initial responses to stereograms lacking a global match has not been studied. In the present study, we addressed this question by examining the relationship between the integration of SF pass-bands and the attenuation of responses to anticorrelated RDSs in monkey extrastriate area V4.

Preliminary results have been reported previously (Kumano et al. 2006).

METHODS

Animal preparation

We studied two adult Japanese monkeys (*Macaca fuscata*) weighing 7 and 8 kg. All animal care, training, and experimental procedures

Address for reprint requests and other correspondence: I. Fujita, Laboratory for Cognitive Neuroscience, Graduate School of Frontier Biosciences, Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka, 560-8531, Japan (E-mail: fujita@fbs.osaka-u.ac.jp).
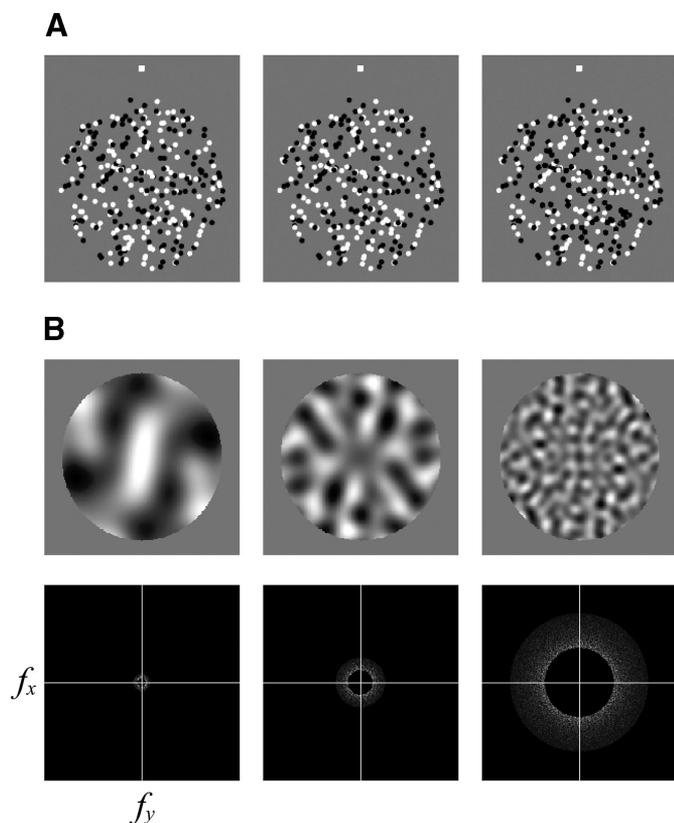
A

B

FIG. 1. Visual stimuli. *A*: random-dot stereogram (RDS) of a center-surround configuration. *Left* and *middle*: images form a stereo pair of a correlated RDS with a small crossed ("near") disparity. Parallel-eyed fusion evokes a perception of a center disk floating over a surrounding annulus. *Right*: image has a center region the luminance contrast of which is inverted relative to the leftmost one. *Middle* and *right*: images thus form an anticorrelated RDS. The binocularly corresponding dots (with an opposite contrast) in the center disk of the RDS generate a small uncrossed ("far") disparity, but a cyclopean surface cannot be perceived. *B*: filtered noise images, each contain a different spatial frequency band (*top*), and corresponding 2-D power spectra (*bottom*). The spatial frequency band increases from left to right. In each spectrum, the origin represents the DC component, and the horizontal and vertical lines denote the spatial frequency in horizontal and vertical directions (f_x and f_y), respectively.

were in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals and were approved by the Animal Experiment Committee of Osaka University.

Each animal underwent aseptic surgeries for the placement of a post for head restraint, a recording chamber, and scleral search coils into both eyes. During the first surgery, we attached a head post and a recording chamber to the skull. The monkey was first premedicated with atropine sulfate (0.03 mg/kg im, Tanabe, Osaka, Japan) to reduce salivation and to promote sedation during surgery and then anesthetized with ketamine HCl (Ketalar, 25 mg/kg im, Sankyo, Tokyo, Japan). Surgical anesthesia was accomplished with sodium pentobarbital (Nembutal, 17 mg/kg iv, Dainippon Sumitomo Pharma, Osaka, Japan). The local anesthetic lidocaine (AstraZeneca Japan, Osaka, Japan) was applied to pressure points or incision sites prior to mounting the monkey in a stereotaxic instrument or making incisions. The top of the skull was exposed, and four holes were drilled through the skull. A head post was attached to the skull and fixed with acrylic resin to four stainless steel bolts inserted into the drilled holes. A custom-fabricated recording chamber was placed over the prelunate gyrus, centered 25 mm dorsal and 5 mm posterior to the ear canals. Throughout each surgery, heart rate was monitored continuously. Body temperature was maintained near 37°C with a heating pad. After surgery, the monkeys were treated with an antibiotic, cefotiam

hydrochloride (Pansporin, 8.0 mg/kg im, Takeda Pharmaceutical, Osaka, Japan), an analgesic, ketoprofen (Menamin, 0.8 mg/kg im, Sanofi Aventis, Tokyo, Japan), and a corticosteroid, dexamethasone sodium phosphate (Decadron, 0.1 mg/kg im, Banyu Pharmaceutical, Tokyo, Japan).

After a recovery period of 2 wk, we performed a second surgery to implant search coils. Administration of drugs and anesthetics were the same as in the first surgery. We implanted a Teflon-insulated stainless steel search coil (Cooner Wire, Chatsworth, CA) in each eye. Subsequently the monkeys were trained to perform a fixation task. After the training was completed, we drilled a hole in the skull to expose the dura mater beneath the recording chamber.

Behavioral task and visual stimulation

The monkeys were seated and head-restrained in a primate chair in a dark, sound-attenuated booth. Here they viewed a visual stimulus on a flat-faced 24-in color display (Sony GDM-FW900, Tokyo, Japan) at a distance of 57 cm. The screen subtended $48 \times 30°$ in visual angle. Stimulus was presented dichoptically by means of a pair of ferro-electric liquid crystal shutters (DisplayTech, Longmont, CO) that were mounted in front of monkey's eyes. Stereo half-images for the left and right eyes were presented alternately at a frame rate of 100 Hz (i.e., 50-Hz refresh for each eye) and were synchronized with refresh of the display. We controlled the behavioral task and data acquisition with a commercial software package (TEMPO, Reflective Computing, St. Louis, MO). Each trial began with the onset of a fixation point. The monkeys were required to move their gaze toward it within 500 ms. Visual stimulus appeared 1 s after the onset of the fixation point for a duration of 1 s. The fixation point dimmed 0.5 s after the offset of the visual stimulus. The fixation window was $2.0 \times 2.0°$. The monkeys were also required to maintain their vergence angle within $\pm0.5°$ of the plane of the screen. They were rewarded with a drop of water for maintaining fixation within these limits for the trial duration. If the monkeys broke fixation during the trial, the trial was terminated, the data were discarded, and the monkeys were not rewarded.

We developed an OpenGL program to generate visual stimuli on the display screen. Gamma correction was applied to ensure a linear relationship between the actual luminance and the gun intensity of the CRT. We used sinusoidal luminance gratings drifting at 2 Hz to examine spatial frequency (SF) tuning of V4 neurons. When neurons did not respond to gratings, we used two-dimensional (2-D) filtered noise images (Fig. 1*B*) to measure spatial frequency tuning. The stereograms for testing the SF tuning were presented binocularly with zero disparity. The 2-D filtered noise images were created by first generating a 2-D binary (i.e., bright or dark) random-noise image and calculating the 2-D Fourier transform of the binary noise, where f_x and f_y are the spatial frequencies along the horizontal and vertical axes, respectively. Then in the Fourier domain spanned by f_x and f_y, a band-pass filter was multiplied with the noise image. The filter was the product of two functions that depended only on the distance from the origin in the Fourier domain. In other words, the two functions depended only on the SF while being uniform over orientation. The two functions were a rectangular window-function and a decaying function that was inversely proportional to the SF, respectively. The bandwidth of the window function was always one octave. The inverse proportionality of the second function provided the pass-bands a fixed energy as long as the window's bandwidth (in octaves of SF) was constant. The filtered image in the Fourier domain was inverse-transformed to the space domain. Ten filtered noise images were iteratively generated before each trial to avoid processing delays interrupting the frame swapping in dynamic filtered noise images refreshed at 10 Hz. Each filtered noise image was computed from an independently generated binary noise.

We measured horizontal disparity tuning of V4 neurons with dynamic RDSs (Fig. 1*A*). Each RDS comprised a central circular disk

and a surrounding annulus (1.0° wide). The center disk was usually tailored for each tested neuron to cover the classical receptive field that we determined manually. If neuronal responses were suppressed by the surround annulus, we reduced the size of the entire RDS until we obtained sufficient responses. The binocular disparity of the center disk was varied from trial to trial, whereas the surrounding annulus was maintained at zero disparity. This ensured no monocular changes in the stimulus associated with varying disparity of the center disk. RDSs were constructed with an equal number of bright (4.6 cd/m^2) and dark (0.3 cd/m^2) dots presented on a mid-level background (2.5 cd/m^2). Each dot was 5×5 pixels ($0.15 \times 0.15°$). The density of dots was 25%. A new dot pattern was generated every 100 ms (10 Hz) with subpixel resolution using anti-aliasing provided by the OpenGL board (Wildcat VP870, 3Dlabs, Milpitas, CA). To minimize stereo cross-talk we used only red phosphors. We did not observe measurable stereo cross-talk with this system employing ferro-electric liquid crystal shutters.

Electrophysiological recording

We used custom-made tungsten-in-glass electrodes or commercially available glass-coated platinum-iridium electrodes (FHC, Bowdoin, ME) with impedance values between 0.3 and 2.0 MΩ (at 1 kHz) for recordings. A pulse motor micromanipulator (MO-951, Narishige, Tokyo, Japan) was mounted onto the recording chamber. The signal from the electrode was amplified and filtered (200–2,000 Hz) with custom-made electronic equipment. We isolated extracellular action potentials or spikes from single neurons with either a custom-made voltage window-discriminator or a spike sorting system based on a template-matching algorithm (Multi Spike Detector; MSD, Alpha-Omega Engineering, Nazareth, Israel). Times of spike occurrences and behavioral events were stored on a computer with 1-ms resolution. We monitored the positions of both eyes with a magnetic search coil system (Enzansi, Tokyo, Japan) and stored it on a disk at a rate of 1 kHz. We identified area V4 by retinotopy, the size-eccentricity relationship of receptive fields, and the positions of the lunate and superior temporal sulci. Histological reconstructions of electrode penetrations in our previous studies have shown that these physiological criteria are reliable indicators for locating V4 (Tanabe et al. 2005; Umeda et al. 2007; Watanabe et al. 2002).

Experimental protocol

After isolating the action potentials of a single neuron, we mapped the position and size of its receptive field using a small patch of either drifting grating or RDS with a zero disparity stimulus. When the neuron responded sufficiently well to a grating patch, we first manually determined its direction preference. Then using a grating that drifted in the neuron's preferred direction, we measured the spatial frequency tuning. If it was unresponsive to gratings, we presented a 2-D filtered noise image instead. We presented spatial frequencies ranging from 0.1 to 6.4 cycle/° (cpd) in octave step. After testing the spatial frequency tuning, we measured binocular disparity tuning with RDSs. For this test, we interleaved both a binocularly correlated RDS (cRDS) and an anticorrelated RDS (aRDS) the binocular disparity of which varied among −1.4, −0.8, −0.4, 0, +0.4, +0.8, +1.4°. Negative and positive values are crossed ("near") and uncrossed ("far") disparities, respectively. We also included a binocularly uncorrelated random-dot pattern, monocular controls (left eye only and right eye only stimulations of random-dots), and a blank condition in our routinely used stimulus set. All stimulus conditions within each repetition were interleaved in pseudo-random order.

Data analysis

The neuronal responses to each stimulus condition were defined as the mean firing rate across trial repetitions over a time window of the visual stimulus duration. The time window of calculating the firing rate was delayed by 80 ms from the visual stimulation period to compensate for the neuronal response latency, which is typically ~100 ms in area V4 (Tanabe et al. 2004). Spontaneous firing rate was calculated from the spike discharges to a blank screen.

We evaluated neuronal tuning curves for spatial frequency and binocular disparity for RDSs by fitting with analytical functions. As the variance of firing rate is proportional to the mean firing rate (Dean 1981; Tolhurst et al. 1981), we minimized this dependence by transforming the firing rate into its square root (Prince et al. 2002). We searched for a parameter combination where the sum of the squared error between the square root of firing rate and the square root of the fitted function reached a minimum.

Spatial frequency tuning curve was fit with a Gaussian function with either a linear or log frequency axis. We adopted the Gaussian function that provided the least error between the two plotting conditions (Read and Cumming 2003). For each neuron, we determined the preferred spatial frequency and the tuning half-width from the best-fitting Gaussian function. The preferred spatial frequency was defined as the frequency at which the fitted curve reached its maximum. The spatial frequency half-width was taken as the ratio in log scale of the preferred spatial frequency and the spatial frequency that produced the half height of the Gaussian amplitude. When the fitted function had two points corresponding to the half-height, we took the mean of the two half-widths.

Disparity-tuning curves were fit with a Gabor function. A Gabor function is the product of a Gaussian and a sinusoid and has six free parameters: the baseline of the curve B, the amplitude A, the center position d_0, and the sigma σ of the Gaussian envelope, and the frequency f and the phase ϕ of the cosine carrier. The disparity-tuning curves for cRDSs and aRDSs were fit simultaneously under the condition that A and ϕ were allowed to vary independently for cRDSs and aRDSs, whereas d_0, B, σ, and f were shared for both cRDSs and aRDSs Gabor functions. The degree to which the tuning amplitude agrees with the prediction of the disparity energy model was quantified as the modulation amplitude ratio: A_a/A_c (the subscripts c and a refer to fitting of cRDSs and aRDSs data, respectively) (Cumming and Parker 1997; Nieder and Wagner 2001; Tanabe et al. 2004).

RESULTS

We recorded from a total of 263 neurons from two monkeys (165 from *monkey H* and 98 from *monkey T*) performing a simple fixation task. Tuning for spatial frequency and for horizontal disparity in both cRDSs and aRDSs were measured from 123 neurons (78 from *monkey H* and 45 from *monkey T*; ≥6 trials with good unit isolation). Seventy-eight neurons of the 123 were selective for binocular disparity in cRDSs (Kruskal-Wallis test, $P < 0.05$). Fifty-two of these 78 (67%) were not disparity selective for aRDSs (Kruskal-Wallis test, $P > 0.05$). This proportion agrees well with our previous report (64%) (Tanabe et al. 2004).

Quantitative description of tuning characteristics

The typical disparity tuning for neurons in V4 was a strong peak at crossed disparity and a shallow dip at uncrossed disparity when the stimulus was a cRDS (Fig. 2A). The disparity tuning was greatly attenuated, or completely lost, when the stimulus was an aRDS (Kruskal-Wallis test, $P < 0.0001$ for cRDSs, $P = 0.57$ for aRDSs). In a small fraction of neurons, the dip in the disparity tuning was as marked as the peak when the stimulus was a cRDS (Fig. 2B). Thus the disparity tuning function was odd symmetric for these neurons. Regardless of this odd symmetry, these neurons also lost
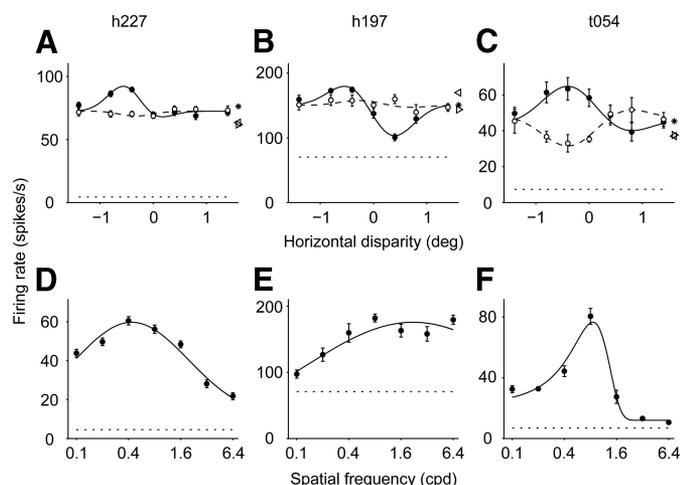
FIG. 2. Horizontal disparity tuning curves and spatial frequency tuning curves for 3 V4 neurons. In each plot, data points and error bars indicate the mean firing rates and SE, respectively. · · · , spontaneous firing rate. *A–C*: horizontal disparity tuning curves for both correlated RDSs (cRDSs, —) and anticorrelated RDSs (aRDSs, - - -). The curves are Gabor functions that provide the best fit. ● and ○: mean responses to cRDSs and aRDSs at each disparity. ◁ and ▷, mean responses of monocular stimulation with random-dots for left eye and right eye, respectively. *, mean response to binocularly uncorrelated random-dots. *D–F*: spatial frequency tuning curves. —, Gaussian function that provides the best fit. *A* and *D*: amplitude ratio (AR) = 0.13, spatial frequency half-width (HW) = 2.17 octaves. *B* and *E*: AR = 0.14, HW = 2.88 octaves. *C* and *F*: AR = 0.81, HW = 1.04 octaves.

disparity tuning when the stimulus was an aRDS. However, the modulation of disparity tuning was not lost in some cases. An example is shown in Fig. 2*C*. For this neuron, the disparity tuning for an aRDS is the inverse of that for a cRDS relative to the baseline response to uncorrelated random-dots (Fig. 2*C, *). The disparity-tuning data recorded in this study reliably reproduced many of the characteristics reported in previous studies (Tanabe et al. 2004, 2005). Thus we conclude that we are adequately obtaining data that reflects the properties of neurons in V4 despite that the sampled neurons were independent and that the steps in tested disparities were different between this and previous studies.

The example two neurons with attenuated disparity tuning for aRDSs displayed a broad bandwidth of SF tuning, regardless of whether the disparity tuning was even or odd symmetric (Fig. 2, *D* and *E*). In contrast, the neuron with similar modulation of disparity tuning for aRDSs and cRDSs had narrow bandwidth of SF tuning (Fig. 2*F*). In these examples, the SF bandwidth was better correlated with the attenuation of disparity tuning between aRDSs and cRDSs than with the shape of the disparity tuning for cRDSs. We looked at whether a similar pattern held across the whole dataset.

We completed the SF tuning test for 236 neurons (≥6 trials), and of these, 233 were selective for spatial frequency (Kruskal-Wallis test, $P < 0.05$). Note that these samples include neurons for which the disparity tuning was not tested. To measure the tuning curves, we fitted a Gaussian function to the SF tuning. The majority of these (62%, 145 of 233) yielded an R^2 value >0.9 with a median of 0.95. The Gabor function was fitted to the disparity tuning for 81 neurons that displayed significant disparity selectivity for either cRDSs or aRDSs (Kruskal-Wallis test, $P < 0.05$). For neurons selective for disparity in cRDSs ($n = 78$), the median R^2 value for cRDSs was 0.83.

Thus these two functions accurately captured the variance in the respective tuning data.

These functions are superimposed on the tuning data in each panel of Fig. 2. The ratio of envelope amplitude for aRDSs to cRDSs was small for the first two representative neurons (0.13 and 0.14, respectively), whereas the ratio was high for the third (0.81). The SF bandwidth, measured as the half-width at half-height, was large for these first two neurons (2.17 and 2.88 octaves, respectively), whereas the bandwidth was small for the third (1.04).

Relationship between attenuation and bandwidth

For analysis of relationship between SF tuning and disparity tuning, we employed several criteria for selecting neurons. First, each neuron was selective for spatial frequency, and for horizontal disparity with either cRDSs or aRDSs (Kruskal-Wallis test, $P < 0.05$). Second, the fitted curve for SF tuning provided an R^2 value >0.6. Third, one or both of the respective R^2 values for the two disparity tunings (cRDSs and aRDSs) exceeded 0.6. Sixty-five neurons met our criterion of $R^2 > 0.6$ for both the disparity tuning and SF tuning (50 from *monkey H* and 15 from *monkey T*). Fifteen neurons were discarded because they did not reach the R^2 criterion, and 1 neuron was discarded because it did not exhibit significant SF selectivity. The lack of SF selectivity indicates an indefinitely broad bandwidth for this neuron. Consistent with the three example neurons in Fig. 2, the amplitude ratio of this neuron was as low as 0.21. Among the 65 neurons included in our analysis, 40 were tested for SF tuning with drifting gratings and 25 were tested with 2-D filtered noise images. The distributions of SF bandwidth did not differ between neurons tested with the two stimuli (Mann-Whitney *U* test, $P = 0.35$).

We obtained both SF bandwidth and amplitude ratio for the 65 neurons. The distributions of SF bandwidth (median: 1.45 octaves) and amplitude ratio (median: 0.38) were similar to those reported in previous studies of V4 neurons (Desimone and Schein 1987 for SF bandwidth, Tanabe et al. 2004 for amplitude ratio). Importantly, these two values were negatively correlated (Spearman's rank correlation $r_s = -0.43$, $P = 0.0004$; Fig. 3*A*). This was true even when the monkey identity was taken in as an independent factor (ANCOVA, $r = -0.13$ for *monkey H*; $r = -0.22$ for *monkey T*; $P < 0.01$ for both monkeys). There was no difference between the two monkeys (ANCOVA, $P = 0.40$). The V4 neurons that are advanced in stereo processing are also the neurons that integrate signals across a wide range of SF channels. This result implies that integration across SF channels contributes to the neural computation of stereo correspondence.

To see if a distinct characteristic of SF tuning exists in the population of disparity selective neurons in V4, we compared the SF half-width between the disparity selective neurons ($n = 73$; $R^2 > 0.6$ for the SF tuning fit) and the disparity nonselective neurons ($n = 45$; $R^2 > 0.6$ for the SF tuning fit). The distributions of the SF half-width were not different among the two populations (Mann-Whitney *U* test, $P = 0.08$). The SF half-width for the entire sample ($n = 203$; $R^2 > 0.6$) ranged from 0.3 to 3.7 octaves with a median of 1.24 octaves. Thus the bandwidth of the neurons included in our main analysis faithfully represents that of the entire population recorded in V4.
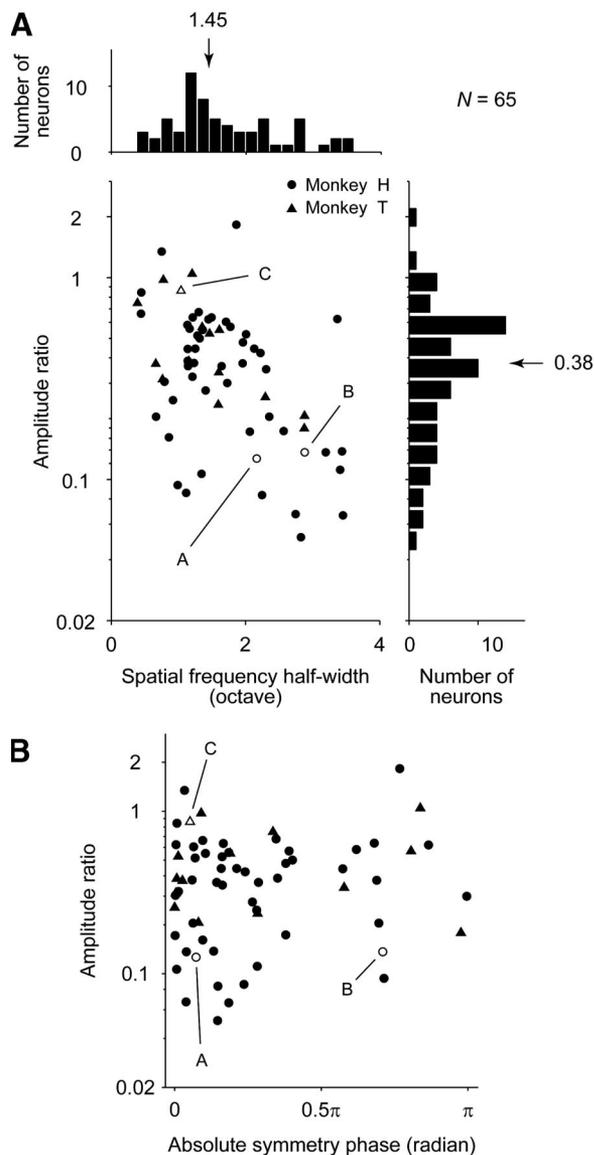
A

B

FIG. 3. *A*: there was a significant negative correlation between spatial frequency half-width and the amplitude ratio (scattergram). The 3 representative neurons in Fig. 2 are denoted by ○ and △. *Top*: distribution of the spatial frequency half-width. The median half-width was 1.45 octaves. *Right*: distribution of the amplitude ratio. The median value of the amplitude ratio was 0.38. *B*: relationship between the amplitude ratio and the shape of the disparity-tuning curve for cRDSs. There was no correlation between these 2 parameters.

The simplest model that can explain the low-amplitude ratio is adding output nonlinearity to the disparity energy model (Lippert and Wagner 2001). For this model to achieve low-amplitude ratio across a variety of disparity-tuning types, tuned-excitatory (TE) neurons need an expansive nonlinearity, whereas tuned-inhibitory (TI) neurons need a compressive nonlinearity. Although there is no evidence to refute this high specificity, we tested this model on the assumption that the nonlinearity generalizes over the population. When the same nonlinearity such as an expansive one is applied to all classes of neurons, the model predicts a systematic relationship between the amplitude ratio and disparity-tuning class; TE neurons show amplitude ratio smaller than unity, TI neurons show amplitude ratio larger than unity, and near and far neurons

show amplitude ratio close to unity (Lippert and Wagner 2001). To see whether this output-nonlinearity model based on a single SF channel sufficiently describes the attenuation, we evaluated the shape of the disparity-tuning curve by calculating the symmetry phase (Read and Cumming 2004) and taking its absolute value. The absolute symmetry phase varies continuously from 0 (TE), $\pi/2$ (near/far) to π (TI). The amplitude ratio did not correlate with the absolute symmetry phase, providing no evidence for this model ($r = 0.02$, $P = 0.23$; Fig. 3*B*). The data suggest that attenuation cannot be accounted for solely by output nonlinearity added to a single SF channel.

DISCUSSION

This study examined the relationship between SF tuning and binocular disparity tuning of neurons in area V4. The amplitude ratio of disparity tuning with cRDSs and aRDSs was the main characteristic of disparity tuning we analyzed. This metric reflects the progression toward a neural representation of the global match solution to the correspondence problem. Neurons with low-amplitude ratio displayed a broad bandwidth of SF tuning. We suggest that integration over multiple SF channels in V4 or at an earlier stage contributes to the attenuation of disparity tuning to anticorrelated stimuli.

Spatial frequency integration for stereo processing

The original algorithm for solving the correspondence problem already encompassed the idea of integrating information across SF channels (Marr and Poggio 1979). The idea was to search for matching patterns iteratively starting from the coarsest channel, then using that match to constrain the search in the next finer channel. This algorithm is incompatible with our current knowledge of the disparity energy computation, where each channel outputs a scalar signal instead of a constraint. Although the stereo computation has drastically been revised from the original proposal, the concept of integrating the outputs of SF channels still holds useful in solving the correspondence problem. A perceptron-type network model that integrates the outputs of multiple disparity-energy subunits can generate flat disparity tuning for aRDSs at the final output (Lippert et al. 2000). The model predicts that if the network acquires such property through learning, the output has a broad SF bandwidth. The integration across disparity-energy subunits is capable of explaining psychophysical performance not only for random-dot patterns but also band-pass noise images (Read 2002).

We learn from these theoretical studies that SF integration is a biologically plausible way of solving the correspondence problem. The models do not give us insights into questions such as where in the visual cortex does integration take place, which range of passbands are integrated, and what is the function that converts the outputs of multiple channels into one scalar value. Neuronal recordings from cats' V1 showed that SF channels are integrated with uneven weights over a range of ~0.3 octave (Menz and Freeman 2003, 2004a,b). The bandwidth of the integration in V1 is too small to account for psychophysically measured integration of ~2 octaves (Farell et al. 2004; Rohaly and Wilson 1993, 1994; Smallman 1995; Wilson et al. 1991). Our study suggests that the integration takes place along the pathway from V1 to V4 and that the integration is completed by the stage of V4.

At the level of the population average, the visual area with a broader SF bandwidth (the median SF full-width at half-height, 2.2–2.5 octaves for V4 and 1.7 octaves for V1) (Desimone and Schein 1987; Foster et al. 1985; present study) has a lower amplitude ratio (the mean amplitude ratio, 0.38 for V4 and 0.52 for V1) (Cumming and Parker 1997; Tanabe et al. 2004). At the level of single neurons in V4, neurons that were less sensitive to disparity in aRDSs had broader SF tuning (Fig. 3*A*). At both levels, the SF bandwidth negatively correlates with the amplitude ratio. These results suggest that integration of a range of SF channels contributes to the attenuation of disparity tuning when the stimulus is devoid of a global match.

Possible alternative explanation

We now consider whether a single disparity energy model with only additional expansive nonlinearity accounts for the negative correlation between amplitude ratio and bandwidth. This model predicts two constraints in the relationship between the passband of the SF tuning, the symmetry phase of the disparity tuning, and the amplitude ratio of the disparity tuning. First, if a neuron's SF tuning is low-pass down to the DC component (infinitely broadband in logarithmic scale), its disparity tuning will have a symmetry phase of zero (or π) and an amplitude ratio smaller than unity. Second, if a neuron's disparity tuning has a symmetry phase of $\pi/2$ (or $-\pi/2$) and an amplitude ratio of unity, its SF tuning is narrowband. These two constraints might give rise to two correlations; a negative correlation between the bandwidth and the amplitude ratio, which is the main result of this paper, as well as positive correlation between the absolute value of the symmetry phase and the amplitude ratio. We did not find the latter correlation in our data (Fig. 3*B*). This suggests that applying static output nonlinearity to the disparity energy model within single SF channel is insufficient for the attenuation.

The second model we consider is a modified version of the preceding one. When a threshold nonlinearity is applied prior to binocular summation of the disparity energy model, the amplitude ratio is lower for models with Gaussian receptive field profiles than for models with Gabor receptive field profiles (Read et al. 2002). This would lead to a negative correlation between SF bandwidth and amplitude ratio. If only V1 neurons with a broad bandwidth and a low-amplitude ratio project to V4, this specific projection pattern might explain the negative correlation we observed (Fig. 3*A*) without invoking SF integration beyond V1. Because no data comparable to those in Fig. 3*A* are available for V1, we should wait for future studies to address this possibility.

Functional significance of broadband neurons

Our results suggest that integration over multiple SF channels contributes to the attenuation of disparity tuning for aRDSs. If there is no normalization in the integration across SF channels, then the more channels are active, the stronger the signal is at the output. In this scenario, a broadband image is an ideal stimulus for generating a strong sensory signal because it activates a broad range of channels. Humans indeed perform better when the image is broadband in some stereo discrimination tasks (Read and Eagle 2000; Westheimer and McKee 1980), suggesting that the integration not only attenuates re-sponses to nonmatching stimuli but facilitates stereo processing when the stimulus has a finite bandwidth.

R E F E R E N C E S

Cogan AI, Lomakin AJ, Rossi AF. Depth in anticorrelated stereograms: effects of spatial density and interocular delay. *Vision Res* 33: 1959–1975, 1993.

Cumming BG, Parker AJ. Responses of primary visual cortical neurons to binocular disparity without depth perception. *Nature* 389: 280–283, 1997.

Cumming BG, Shapiro SE, Parker AJ. Disparity detection in anticorrelated stereograms. *Perception* 27: 1367–1377, 1998.

Dean AF. The variability of discharge of simple cells in the cat striate cortex. *Exp Brain Res* 44: 437–440, 1981.

Desimone R, Schein SJ. Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *J Neurophysiol* 57: 835–868, 1987.

Farell B, Li S, McKee SP. Coarse scales, fine scales, and their interactions in stereo vision. *J Vis* 4: 488–499, 2004.

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

Foster KH, Gaska JP, Nagler M, Pollen DA. Spatial and temporal frequency selectivity of neurones in visual cortical areas V1 and V2 of the macaque monkey. *J Physiol* 365: 331–363, 1985.

Janssen P, Vogels R, Liu Y, Orban GA. At least at the level of inferior temporal cortex, the stereo correspondence problem is solved. *Neuron* 37: 693–701, 2003.

Julesz B. *Foundations of Cyclopean Perception.* Chicago, IL: University of Chicago, 1971.

Kumano H, Tanabe S, Fujita I. Spatial frequency integration for stereo processing in macaque visual area V4 (abstract). *J Vision* 6: 895a, 2006.

Lippert J, Fleet DJ, Wagner H. Disparity tuning as simulated by a neural net. *Biol Cybern* 83: 61–72, 2000.

Lippert J, Wagner H. A threshold explains modulation of neural responses to opposite-contrast stereograms. *Neuroreport* 12: 3205–3208, 2001.

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

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

Menz MD, Freeman RD. Functional connectivity of disparity-tuned neurons in the visual cortex. *J Neurophysiol* 91: 1794–1807, 2004a.

Menz MD, Freeman RD. Temporal dynamics of binocular disparity processing in the central visual pathway. *J Neurophysiol* 91: 1782–1793, 2004b.

Nieder A, Wagner H. Hierarchical processing of horizontal disparity information in the visual forebrain of behaving owls. *J Neurosci* 21: 4514–4522, 2001.

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

Prince SJD, Pointon AD, Cumming BG, Parker AJ. Quantitative analysis of the responses of V1 neurons to horizontal disparity in dynamic random-dot stereograms. *J Neurophysiol* 87: 191–208, 2002.

Read JCA. A Bayesian model of stereopsis depth and motion direction discrimination. *Biol Cybern* 86: 117–136, 2002.

Read JCA, Cumming BG. Testing quantitative models of binocular disparity selectivity in primary visual cortex. *J Neurophysiol* 90: 2795–2817, 2003.

Read JCA, Cumming BG. Ocular dominance predicts neither strength nor class of disparity selectivity with random-dot stimuli in primate V1. *J Neurophysiol* 91: 1271–1281, 2004.

Read JCA, Eagle RA. Reversed stereo depth and motion direction with anti-correlated stimuli. *Vision Res* 40: 3345–3358, 2000.

Read JCA, Parker AJ, Cumming BG. A simple model accounts for the response of disparity-tuned V1 neurons to anticorrelated images. *Vis Neurosci* 19: 735–753, 2002.

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

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

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

Tanabe S, Doi T, Umeda K, Fujita I. Disparity-tuning characteristics of neuronal responses to dynamic random-dot stereograms in macaque visual area V4. *J Neurophysiol* 94: 2683–2699, 2005.

Tanabe S, Umeda K, Fujita I. Rejection of false matches for binocular correspondence in macaque visual cortical area V4. *J Neurosci* 24: 8170–8180, 2004.

Tolhurst DJ, Movshon JA, Thompson ID. The dependence of response amplitude and variance of cat visual cortical neurones on stimulus contrast. *Exp Brain Res* 41: 414–419, 1981.

Umeda K, Tanabe S, Fujita I. Representation of stereoscopic depth based on relative disparity in macaque area V4. *J Neurophysiol* 98: 241–252, 2007.

Watanabe M, Tanaka H, Uka T, Fujita I. Disparity-selective neurons in area V4 of macaque monkeys. *J Neurophysiol* 87: 1960–1973, 2002.

Westheimer G, McKee SP. Stereoscopic acuity with defocused and spatially filtered retinal images. *J Opt Soc Am* 70: 772–778, 1980.

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