Binocular spatial frequency profiles of disparity-sensitive neurons in the primary visual cortex

初期視覚野視差選択性細胞における両眼空間周波数特性

馬場美香

平成 27 年 9 月修了

ABSTRACT

For our vivid perception of a three-dimensional world, the stereoscopic function begins in our brain by detecting slight shifts of image features between the two eyes, called 'binocular disparity'. The primary visual cortex is the first stage of this processing, and neurons there are tuned to a limited range of spatial frequencies (SF). However, our visual world is generally highly complex, composed of numerous features at a variety of scales, thereby having broadband SF spectra. This means that binocular information signaled by individual neurons is incomplete, and combining information across multiple SF bands must be essential for the visual system to function in a robust and reliable manner. In this study, I investigated whether the integration of information from multiple SF channels begins in the primary visual cortex. I measured disparity-selective responses in the joint left-right SF domain using sequences of dichoptically flashed grating stimuli consisting of various combinations of SFs and phases. The obtained interaction map in the joint left-right SF domain reflects the degree of integration across different SF channels. A substantial fraction of complex cells in my data showed highly elongated binocular SF profile, which is consistent with the idea that disparity information is combined from multiple SF channels. For the majority of these neurons, the optimal disparity is matched across the SF bands. In addition, some of the complex cells showed extremely sharp SF tuning, i.e., narrow bandwidth for binocular SF matching for disparity detection, compared with relatively broad SF bandwidth for monocular response. It suggests that both narrowband binocular SF matching and integration of broadband SF information are achieved simultaneously for accurate disparity detection. These results suggest that a highly specific SF integration process for disparity detection starts in the primary visual cortex.

TABLE OF CONTENTS

GENERAL INTRODUCTION 5
Wavelet analysis of retinal images in the primary visual cortex 5
Binocular disparity and its detection 7
Disparity energy model 8
INTRODUCTION 10
MATERIALS AND METHODS 13
Animal preparation and maintenance 13
Visual stimulation 14
Data Analysis 15
RESULTS 17
Selectivity to binocular phase combinations for various SF pairs between the two eyes 17
Substantially elongated binocular SF interaction maps for a subset of complex cells 19
Reconstruction of binocular receptive fields (bRFs) and disparity tunings in the space domain 22
Are optimal disparities matched across different SF bands? 23
Relationship between pooling in the SF domain and the space domain 25
Binocular matching with substantially sharp SF tunings 27
DISCUSSION 29
Interactions of pooling across multiple stimulus dimensions 29
Binocular matching is conducted with substantially sharp SF tunings 29
Relation to coarse-to-fine mechanism 30
Possible effects of suppressive elements 31
Alignment of disparity tuning curves for pooling across SF bands 32
Effects of SF pooling on the shape of disparity tuning curve 33
REFERENCES 34

Acknowledgements

I really thank Prof. Izumi Ohzawa for his strict but at the same time, kind and tolerant guides for all of things throughout the period of my graduate student life. I was given a lot of important way of thinking from him, not only about research or education but also about a view of life. I cannot thank peoples in Ohzawa lab enough for their kind advices, strict and appropriate criticisms, and great supports. Especially, I wish to thank Dr. K.S. Sasaki for his educations and great supports at anytime throughout my student life and my thesis. I was always taught about mathematics, physiology, experiment, logical thinking and even, English from him. I wish to thank D. Kato for many discussions, rigid and appropriate criticisms for all of my works, and kind help, especially in mathematics and programming. I thank T. Nakazono for discussions, many aspects in programming, logics, and many things for the experiment systems. I wish to thank K. Okada for his kind discussions and many helps. I wish to thank all of the current lab members for the help in experiments and discussions, and giving me joyful lab life. I thank Y. Kobayashi for giving me a lot of knowledge. I really thank previous lab members, M. Inagaki, H. Tanaka, Y. Asada, T. Arai, S. Nishimoto, T.M. Sanada, T. Ninomiya, and M. Fukui for help in experiments and discussions, experiment setup, and their research issues. I wish to thank T. Tanaka and S. Fujioka for kind supports of my everyday life in lab.

Finally, I really thank my parents for their active and great supports in both financial and mental aspects.

I thank the supports by Ministry of Education, Culture, Sports, Science and Technology Grants KAKENHI 22135006 and 24700325.

GENERAL INTRODUCTION

We can recognize various complex features in our visual world such as faces, objects, textures, surface qualities of materials or particular movements, and three-dimensional depth, very easily without any special effort. For such outstanding ability, visual system in our brain possesses highly specialized function of extracting various kinds of visual features from our retinal images (Gross et al., 1972; Orban G. A, 1995; Tanaka K, 1996; Cumming and DeAngelis, 2001; Tsao et al., 2006; Nishio et al, 2012). The purpose of this research, from a very general and broad viewpoint, is to understand mechanism at the intermediate-stage of such specialized feature detection, in particular, binocular vision related to the stereopsis. I will start this paper by introducing the first stage of visual processing in our brain.

Wavelet analysis of retinal images in the primary visual cortex

Visual information captured by retinae is processed hierarchically after transmitted to our brain. For an initial stage of retinal image processing in the cerebral cortex, there is an area named 'primary visual cortex (A17)'. Each neuron in A17 receives visual inputs only within a specific small area in the visual field, ranging approximately 1 to 5 degrees of visual angle in diameter (Gattass et al., 1987; Freeman and Simoncelli, 2011). This small area is called a '***receptive field (RF)***'. Different neurons have RFs at different positions in our retina and the whole neuronal population covers the entire visual field. The characteristic property of these neurons in A17 is their strict selectivity to the image pattern, that is, selectivity to a particular 'wave' or 'stripe pattern'. Each neuron strongly responds to a small piece of wave pattern with specific orientation and fineness at its RF position. Since a single piece of wave is defined by its orientation, spatial frequency (SF) and position, the functional property of a neuron in A17 is well described by its selectivity to orientation, SF and RF position.

Since the beginning of A17 studies, neurons in A17 have been classified into two types (Hubel and Wiesel, 1962; Movshon et al., 1978a,b). One of the two types, '***simple cells***', have separated areas (so-called ON region and OFF region) within its RF. ON region is an area responding to bright stimuli and OFF region is an area responding to dark stimuli. Figure 1A shows various RFs of simple cells arranged according to the

orientation and SF, measured in cat A17. For a simple cell, stripe pattern or a piece of wave, does not always cause its response even though the pattern has its optimal orientation and SF. In addition to these stimulus parameters, the position or 'phase' of the wave within the RF influence the activity of a simple cell. Compared with this, the other type called '***complex cells***' generally has no such separated sub-regions within its RFs while they are also highly selective to the orientation and SF of the stimulus. Complex cells are activated whenever a wave with optimal SF and orientation is presented within their RF regardless of its absolute position.

In any case, both simple cells and complex cells possess a strict selectivity to pieces of wave patterns. A piece of wave is called as a '***wavelet***', and the processing of retinal image in A17 can be understood from a viewpoint of ***'wavelet analysis'***. Especially, shapes of RFs in A17 are well described by 'Gabor function' (Fig. 1B), and called 'Gabor wavelet'. In wavelet analysis, an image is decomposed into many pieces of wavelets, and information of luminance pattern in the x-y space is represented as a combination of strengths of many different wavelets. A single neuron in A17 works as a channel which detects a particular Gabor wavelet by having selectivity to the position, SF, and orientation of a stripe pattern. Only a specific SF and orientation component located at a particular retinal position can elicit a response in a specific channel. For example, as shown in Figure 1C, a horizontal very fine visual feature, such as the frets of a guitar, is detected by a channel which has the small RF of a Gabor wavelet with horizontal orientation and high SF, while a vertical coarse feature (e.g., neck of the guitar) is detected by that with vertical orientation and low SF. And any arbitrary image can be represented by a complete set of such channels (Lee, 1996). Viewing visual information in the Gabor wavelet analysis is particularly important for studying visual processing in our brain.

Wavelet analysis is a general and very strong technique in image processing applicable to any kind of images. However, decomposed information of a single channel itself usually doesn't have enough information for our higher-order recognition. Natural features we recognize or utilize in our everyday life are generally much more complex, compared with stripe patterns (e.g., faces, textures and so on). To obtain various kinds of feature selectivity, information of a single wavelet is incomplete, and many of such different wavelets should be combined and refined in our visual pathway eventually. Since different visual features are characterized as different combinations of

information in the wavelet pyramid, it may be helpful to consider what is needed for extracting a specific feature in the wavelet pyramid, or in the orientation domain and the SF domain. For example, 'velocity' is one of the visual features represented in area MT in the cortex (Movshon et al., 2986; Maunsell & Newsome, 1987). Although I didn't introduce temporal selectivity of A17 neurons up to here, many of them have optimal temporal frequency (TF) and direction of motion (Hubel & Weisel, 1959; DeAngelis et al., 1993a,b). A particular velocity is characterized by a particular combination of A17 spatio-temporal frequency channels (Simocelli and Heeger, 1992), and neural selectivity in MT may be constructed by combining outputs of them.

In this study, I focus on one visual feature, which is defined in the 'binocular' condition. Before introducing the specific purpose of my study, I will explain the details of that binocular feature, and then explain the current understanding of mechanisms of its detection.

Binocular disparity and its detection

We can perceive a vivid three-dimensional world although our retinal images are two-dimensional. Various monocular and binocular features contained in our retinal images are available for estimating depth, and our brain can compute the three-dimensional arrangement of what we see based on such cues. Especially, the function of estimating depth based on binocular cues is called '***stereopsis***'. Neural mechanisms for stereopsis have been actively pursued with psychophysical (Hemholtz, 1909; Westheimer and McKee, 1980; Schor et al., 1983), physiological (Barlow et al., 1967; Poggio and Fischer, 1977; Ferster 1981; Maunsell and Van Essen, 1983; Ohzawa et al., 1990; Uka et al., 2000; Cumming and DeAngelis, 2001; Tanabe et al., 2011), and computational (Marr and Poggio, 1979; Qian, 1994; Doi and Fujita, 2014; Li and Qian, 2015) approaches.

Because our left and right eyes are located at slightly different positions, left and right images projected on our retinae are very similar, but also slightly shifted by a small distance between the two eyes. This small difference or shift between the two retinal images is called '***binocular disparity***'. The magnitude of binocular disparity is systematically related to three-dimensional depth due to geometrical relationship. Visual inputs at different depth are projected on left and right retinae with different

amounts of offset between the two eyes (Fig. 2A). If we can measure binocular disparity of an object at a certain position accurately, we can basically estimate how far that object is along the depth direction from the fixated plane (e.g., green curve in Fig. 2A).

It is well known that, the detection process of such binocular disparity starts in A17 in our brain. A substantial number of neurons in A17 have their RFs in both left and right eyes and are called 'binocular' neurons (Hubel and Wiesel, 1962). Among them, there are neurons that are especially selective to binocular disparity. Here, since binocular neurons in A17 has optimal orientation and that orientation is generally common across the left and right eyes, the disparity that a single neuron can encode is basically along the direction orthogonal to its optimal orientation (Fig. 2B). As described in the previous section, a simple cell is selective to the position of its optimal wave pattern within its RF. Therefore, a binocular simple cell works partially as a disparity sensor by having the best positional arrangement of the stimulus combination between the two eyes to generate response. However, disparity is not the only parameter which affects the firing rate of a simple cell. Other stimulus parameters, such as absolute monocular position and contrast, also affect the firing rate of it.

On the other hand, nearly a half of binocular complex cells in A17 are also selective to binocular disparity (Ohzawa and Freeman, 1986b), and their property is more suitable as a specialized detector for disparity than simple cells. An ideal disparity detector should respond to particular disparity regardless of other irrelevant visual features (i.e., irrelevant for disparity sensing but may be useful for other purposes). Disparity-sensitive complex cells in A17 basically satisfy such requirements (Ohzawa et al., 1990). For example, they respond strongly to a pair of bars with the same contrast when they are presented dichoptically at a particular positional relationship between the two eyes, anywhere within their RFs (i.e., regardless of absolute monocular stimulus positions). They respond to both conditions of bright-bright pair and dark-dark pair, as long as relative position between the two eyes is retained. The fundamental response properties of such disparity-sensitive complex cells are well explained by the '***disparity energy model,***' which was proposed by Ohzawa et al (1990).

Disparity energy model

The disparity energy model is constructed by combining several wavelets under special constraints so as to produce selectivity to binocular disparity while removing undesired selectivity to other irrelevant parameters as noted above. The model integrates outputs from four binocular simple-type subunits. A whole structure of the model is illustrated schematically in Figure 3. Each subunit (illustrated as a columns in the upper part of Fig. 3) has RFs in both left and right eyes where their relative positions have an offset by a particular distance horizontally, along the direction orthogonal to its optimal orientation (called as 'Position model'). The SF and orientation of subunits are common for both eyes and for four subunits. A critical point in the model is that the four subunits are in *quadrature phase,* the phase of RFs differs from that of its counterpart by 90°. This allows the model to lose its monocular phase selectivity (a hallmark of a complex cell), making the cell responsive to a stimulus anywhere within the RF, while retaining the disparity selectivity. Outputs from all the subunits are half-wave rectified, squared and combined in a single complex cell. Composed as such structure, the model successfully obtains basic property needed for an ideal disparity detector, and explains well many aspects of responses of actual disparity-sensitive complex cells (Ohzawa et al., 1990).

The disparity energy model, however, has still some aspects that are not suitable for ideal disparity detection, known as responses to false-disparities (Cumming and Parker, 1997). Further processing is needed in the next stage as described in the following chapter. In this study therefore, I wish to consider and examine disparity specific processing that goes beyond the conventional disparity energy model, constructed by combining information from wavelets in different scales.

INTRODUCTION

As noted in General Introduction, our visual system possesses a remarkable function called stereopsis that allows perception of three-dimensional depths based on a pair of two-dimensional retinal images. As shown in a pair of stereo photographs in Figure 4A and their cross-sections in Figure 4B, left and right images are generally similar but slightly shifted versions of each other due to the lateral placement of the eyes. Accurate measurement of binocular disparity, that is, such small shifts between left and right retinal images, is a central problem of stereopsis. Here, I try to examine the selectivity of a neuron which is specialized for detection of binocular disparity in the SF domain. I especially focus on the SF domain for the following two reasons. One is that SF is one of the major parameters in Gabor wavelet analysis where the visual information of retinal images is initially processed in the cortex. The other is that, for a long time historically, combining signals of different SF channels has been pointed out to be quite important for improvement of detection accuracy of binocular disparity in computational studies (Marr and Poggio, 1979; Quam, 1987; Fleet et al., 1996; Chen and Qian, 2004).

Generally, our external world is highly complex, with features at many scales existing simultaneously side-by-side or with overlap, sometimes with transparency, texture, shadows and sharp edges of luminance. Clearly therefore, natural scenes are broad-band, containing a broad range of SF components as shown in the spectral distributions in Figure 4C. Therefore, for a robust detection of binocular disparity, the brain must be able to utilize such broadband information. Since each neuron in A17 is tuned to relatively narrow range in the SF domain (red filled curves in Fig. 4C) having a bandwidth of approximately 1.3 octave on average (Movshon et al., 1978), integrating signals from different SF channels may take place at some stage on the disparity processing pathway.

Note that, all the SF channels that are to be combined must be carefully selected under the specific rule to achieve accurate disparity detection. If not, the integrated profile would lose precise selectivity to binocular disparity feature. What combinations of SF information should elicit the responses of disparity-detector neurons? For the simplicity of explanation, from this point on, I will consider the disparity for one-dimensional luminance patterns along the direction orthogonal to optimal

orientation of each A17 neuron. Since any complex pattern can be generated as a sum of various sine-waves of different SFs with appropriate phase and amplitude (Fourier transform), it is convenient and sufficient to consider such various sine-waves (or SF components) instead of a complex luminance pattern to discuss binocular disparity in the SF domain (Figure 5). When luminance patterns are shifted, all of the SF components contained in the patterns are shifted by the same amount across different SFs. Since binocular disparity is basically equal to the size of the shift of patterns between the two eyes as shown in Figure 5A, this constraint is characterized in the SF domain as a shift of sine-waves between the two eyes, by a constant amount in position across all of different SFs (Fig. 5B).

Taking these considerations into account, neurons that achieve accurate disparity detection should be selective to the same inter-ocular position shifts across broad SF band. Such neurons may be constructed by integrating or pooling outputs from multiple disparity energy units, tuned to different SFs but to common disparity. To examine whether such specialized processing happens, measuring binocular SF profiles as follows is applicable. For a neuron consisting only of a single SF channel, predicted disparity selective response would be a circular region of joint left and right one-dimensional SF tunings, in the joint left-right SF domain (Fig. 6A). And because most A17 neurons show nearly identical SF tuning for the two eyes, the circular binocular region would be centered on 1:1 diagonal. On the other hand, for a neuron that integrates or pools multiple of SF channels each tuned to different SFs, the combined response would be line up on the 1:1 diagonal as shown in Figure 6B. For such neurons, a response profile would be an elongated region along 45° diagonal in the joint left-right SF domain. At the same time, they would show common optimal disparity across different SF channels.

Do disparity-sensitive neurons in A17 show such binocular SF profiles? Although a previous study reported a tendency for broadly tuned V4 neurons to show weak responses to anti-correlated (contrast inverted) stereograms (Kumano et al., 2008), no direct measurement of binocular SF profiles described above has been attempted to date. In this study, I first examine whether binocular SF profiles are elongated along the 45° diagonal direction for disparity-sensitive neurons of A17 of the cat, by measuring binocular interactions in the joint left-right SF domain. Next, I examine the consistency

of disparity preference across different SF bands to demonstrate that binocular SF information is precisely integrated to achieve accurate disparity detection.

MATERIALS AND METHODS

Extracellular single-unit recordings were performed in area 17 of 23 anesthetized and paralyzed adult cats (14 males and 9 females, 2.0 kg to 4.5kg). Details about surgical procedure, animal maintenance, and single-unit recording are described previously (Sasaki and Ohzawa, 2007; Ninomiya et al., 2012). Only a brief account of the basic procedures and points different from the previous studies are provided here. All animal care and experimental procedures conformed to the guidelines established by the National Institutes of Health and were approved by the Osaka University Animal Care and Use Committee.

Animal preparation and maintenance

After initial pre-anesthetic doses of hydroxyzine (Atarax, 2.5 mg) and atropine (0.05 mg), anesthesia was induced and maintained with isoflurane (2–3.5% in O_2) for the remainder of the surgical preparation. During surgery, lidocaine was injected subcutaneously or applied topically at all points of pressure and possible sources of pain. A body temperature was monitored and maintained near 38℃ with a servo-controlled heating pad, and ECG electrodes were placed for monitoring heart rate. A tracheostomy was performed for the subsequent artificial respiration. After the animal was secured in a stereotaxic apparatus, anesthesia was switched to sodium thiopental (Ravonal, 1.0 mg・kg^{-1}・h^{-1}) and paralysis was induced with an initial dose of gallamine triethiodide (Flaxedil, 10 mg・kg^{-1}・h^{-1}). Artificial ventilation was performed with a gas mixture of 70% N_2O and 30% O_2. The respiration rate and stroke volume were adjusted to maintain the end-tidal CO_2 between 3.5 and 4.3% throughout the experiment. A craniotomy was then performed over the central representation of the visual field of area 17 approximately at Horsley-Clarke coordinates P4 and L2. Pupils were dilated with 1% atropine sulfate, and nictitating membranes were retracted with 5% phenylephrine hydrochloride (Neosynesin). The corneas were protected using contact lenses of appropriate power with a 3- to 4-mm artificial pupil.

To record single unit activities, tungsten electrodes (A-M Systems) were lowered into a region of cortex exposed with craniotomy. Agar was applied around the electrodes to prevent drying, and melted wax was layered over the agar to seal as a chamber and

reduce cortical pulsation. Electrical signals from the electrodes were amplified (x10,000) and band-pass filtered (300–5,000 Hz). Spikes were sorted by their waveforms and time-stamped with 40 μs resolution (Ohzawa et al., 1996). When the electrodes were retracted, electrolytic lesions were made at intervals of 500–1,200 μm for each electrode track.

At the end of an experiment, the animal was administered an overdose of pentobarbital sodium (Nembutal), perfused with buffered saline solution followed by 4% formalin in 0.1M phosphate-buffered saline (PBS), and cortical tissue was prepared for histological examination. Electrode tracks were reconstructed and cortical laminae were identified.

Visual stimulation

Visual stimuli were generated by computer and displayed on a cathode ray tube display (a resolution of 1,600 x 1,024 pixels, refreshed at 76 Hz; GDM-FW900, Sony) using only green channel to avoid color misconvergence across channels. In each experiment, the luminance nonlinearity of the display was measured using a photometer (Minolta CS-100) and linearized by gamma-corrected lookup tables. 50% Michelson contrast was used for all grating stimuli used in this study. A haploscope was used to present stimuli to the left and right eye separately, and the visual fields of the monitor covered approximately 23° x30° for each eye at a viewing distance 57cm.

Once a single unit was isolated, its preferred orientation, SF, the center position and the size of a receptive field (RF) were tested preliminarily under manual control. Subsequently, sub-space mapping to measure tuning in the joint orientation and SF domain was performed for each eye with flashed gratings (Ringach et al., 1997; Nishimoto et al., 2005). After the optimal orientation, optimal SF and the range of SF were determined for each eye, the binocular measurement was conducted to obtain a binocular SF interaction profile in a 4-dimensional domain (sf_L, sf_R, ph_L, ph_R), where "sf" and "ph" denote SF and phase, respectively, and the subscript indicates the eye. The stimuli were flashed gratings of various combinations of SFs and phases between the two eyes, oriented at the optimal orientation of the target cell (Ninomiya et al., 2012). Basically, 12 SFs and 8 phases per each eye were used, so the total number of stimuli presented in a single block was 9472, i.e., 9216 binocular (12x12x8x8), 192 monocular (2 eyes x12x8), plus blank (8) conditions. The multiple blank conditions were used to

increase the reliability of base-line response estimation. The blank stimuli had the same uniform luminance for both left and right display areas. All 9472 conditions were presented in one randomized block, and the block was repeated about 8 to 20 times. The stimuli were presented in a randomized sequence updated at 38Hz, and its size was adjusted to be slightly larger (approximately 1.5–3 times) than the size of the RF. The range of the SF was set to cover the cell's SF band sufficiently. After the binocular 4-dimensional mapping was completed, tuning curves for orientation and SF were verified respectively using drifting grating stimuli. In some cases, the RFs of the cells were also measured using a standard reverse correlation procedure with dense white noise stimuli (Nishimoto et al., 2006; Sasaki and Ohzawa, 2007).

Data Analysis

Each cell was classified into simple or complex based on standard criteria (F1/F0 ratio) (Skottun et al., 1991). The balance of responses between the two eyes was quantified using the binocularity index (Sasaki and Ohzawa, 2010).

In this study, a goal is to elucidate a specific SF integration for disparity detection as illustrated in Figure 6 where neural responses to various left- and right SF combinations are examined. The actual experiment requires varing phases of the grating stimuli for each eye, because A17 neurons are highly sensitive to variations of stimulus phase monocularly (DeValois et al., 1978; Movshon et al., 1978a) for simple cells, and to variations of inter-ocular phases binocularly (Freeman and Robson, 1982; Ohzawa et al., 1986b) for both simple and complex cells. This means that such an experiment must be done in the 4-dimensional stimulus space (sf_L, sf_R, ph_L, ph_R). Such an experiment will completely specify the binocular disparity selectivity for all possible combinations of SFs and phases for the left and right eye stimuli.

The 4-dimensional data set may be examined in various ways, but for my purposes, it is necessary to reduce the 4-d data into the 2-d form depicted in Figure 6. This process was conducted by following 3 steps as shown in Figure 7. First, for each combination of (sf_L, sf_R) (a square in Fig. 7B), I obtained a binocular phase combination selectivity. This created a map in the joint (ph_L, ph_R) domain (Fig. 7C), and for a disparity energy unit (Ohzawa et al., 1990) for example, it has a form depicted in Figure 7D. Second, I then computed an inter-ocular phase tuning curve (Fig. 7E) by integrating the map

along constant inter-ocular phase lines in Figure 7D (e.g. red straight and broken lines). The final step was to extract the strength of the tuning, which was given by the amplitude of a one-cycle sinusoid fitted to the inter-ocular phase tuning curve. The amplitude was obtained via Fourier analysis. Repeating these steps for all (sf_L, sf_R) combinations gave me a desired map as shown in Figure 6.

A complete data set in the 4-dimensional stimulus space (sf_L, sf_R, ph_L, ph_R) was constructed via a standard spike-triggered averaging (reverse correlation) as depicted in Figure 7A. Such a binocular interaction profile (binocular SF interaction map) was calculated for every recorded cell and evaluated. Specific analyses are described at the relevant places in Results.

RESULTS

I recorded from a total of 74 cells in area 17 of adult cats. Minimum set of measurements could be performed on 18 simple cells, 49 complex cells and 7 unclassified type. For some specific types of measurements, the number of cells was reduced.

Selectivity to binocular phase combinations for various SF pairs between the two eyes

As noted in Introduction, specific integration or pooling of binocular SF channels must happen at some stage along the visual pathway for robust estimation of binocular disparity. Examining this is achieved by measuring disparity selective responses for a variety of left-right combinations of different SFs. Hereafter, I call the disparity-selective profile in the joint left-right SF domain (Figure 6) as a 'binocular SF interaction map'. To quantify the response strength to be plotted in the binocular SF interaction map, I computed the amplitude of disparity tuning for each combination of left-right SFs (sf_L, sf_R). This is because I am interested in the mechanism by which disparity specific information is combined, and not those that are due to monocular excitations. For a given combination of (sf_L, sf_R), I have a complete set of response data for all combination of different left and right phases (ph_L, ph_R). This gives me the phase-based disparity tuning for each (sf_L, sf_R), from which the amplitude of disparity tuning is computed easily. Therefore, once I obtain response data (spike counts) into bins defined in the 4-dimensional parameter space (sf_L, sf_R, ph_L, ph_R), the binocular SF interaction map as well as disparity tunings for all combinations of (sf_L, sf_R) can be computed.

The construction of the complete histograms in the original (sf_L, sf_R, ph_L, ph_R) space was carried out via reverse correlation as illustrated in Figure 7. The responses of neurons were measured while sine-wave grating pairs were presented dichoptically for various combinations of SFs and phases. Stimuli were delivered in rapid flashes at 38 stimuli/sec. Since it is impossible to graphically present histograms in a 4-dimensional space, I present our data in the format depicted in Figure 7B and C: as a matrix of left-right phase tuning profiles (ph_L, ph_R). Here, (ph_L, ph_R) map is shown in Figure 7C

and many of these are arranged into a matrix of (sf_L, sf_R) in Figure 7B. Each binocular phase-domain map generally take a map as shown in Figure 7D for a complex cell that is a disparity energy unit. Data were further reduced to obtain a disparity tuning curve as a function of inter-ocular phase difference (Fig. 7E). The amplitude of modulation in such a disparity tuning curve is a good metric of the degree of binocular interaction for a given (sf_L, sf_R).

Figure 8A presents the data from a representative complex cell in the format described in Figure 7B and C. Each small map shows responses to binocular combinations of phases (ph_L, ph_R) for one SF pair (sf_L, sf_R), and such maps are arranged as a matrix of 13×13 binocular SF pairs. One of the maps marked with red border (sf_L: 0.48 cpd, sf_R: 0.57 cpd) is magnified in Figure 8B to show typical response property to phase combinations of the disparity selective complex cell. Clear response of band in 45° diagonals indicates that this cell has selectivity to a particular phase difference between the two eyes, in this case, approximately 0°. Note that the interpretation of 0° phase difference requires care since the animal was in a paralyzed and anesthetized state without active fixation.

To create a binocular SF interaction map for this neuron, I next evaluated the strength of disparity tuning for each (sf_L, sf_R) combination. This was done by integrating the response profile of Figure 8B along 45° paths (constant inter-ocular phase difference lines) to obtain a one-dimensional tuning curve as a function of inter-ocular phase difference (Fig. 8C). A degree of modulation of this tuning curve was extracted by fitting a sinusoid for the analysis described later (green broken curve). In Figure 8D, such tuning curves are shown for all (sf_L, sf_R) combinations as a matrix similar to Figure 8A in its arrangement. Notice that tuning curve has a strong modulation only for relatively matched SF pairs between the two eyes, and such modulation cannot be observed if difference of SF between the two eyes is large. The evidence of elongation is already clear visually in Figure 8A and 8C. However, one of the most likely candidates for artifacts is a contamination from multiple neurons each tuned to different SFs. In order to reduce this possibility, I have examined recorded spike waveforms for signs of multiple spike waveforms as shown in inset of Figure 8A. Although such an examination of spike waveforms does not completely rule out the possibility of multi-spike contamination, it provides a reasonable safeguard for this artifact. Therefore,

spike waveforms were examined carefully for all neurons identified to have substantial elongations.

Results from a representative simple cell are shown in Figure 8E-H. For this simple cell, response to (ph_L, ph_R) combinations shows a single peak centered at a particular phase pairs, rather than spreading along diagonal band (Fig. 8A and B). This result is obvious because a simple cell has a selectivity to the phase presented monocularly, and the intersection of the two monocular peaks becomes the peak in the joint domain.

Substantially elongated binocular SF interaction maps for a subset of complex cells

To visualize the degree of modulation in Figure 8D and H clearly, maps were created as density plots where each pixel represents the degree of modulation at each (sf_L, sf_R). Values were obtained as the amplitude of the first harmonic (F1) component of an inter-ocular phase tuning curve. Figure 9A presents such a plot derived from the complex cell data in Figure 8D. To evaluate the degree of elongation of the profile as I discussed in Figure 6, the map was fitted with a 2-dimensional Gaussian function whose axes were constrained to the direction of 45° and 135° diagonals (Fig. 9B). The fitted function was always a Gaussian. In some cases, data were measured in logarithmically constant steps. In such cases, data and a contour of the fitted Gaussian were plotted in a logarithmic domain, resulting in the egg-shaped distortion. Elongation index in the SF domain (EIsf) was then computed as a ratio of a fitted sigma along the 45° axis to that along the 135° axis. Clearly, with EIsf = 3.13, this complex cell exhibits a highly elongated shape along the 45° diagonal.

Additional binocular SF interaction maps are shown in Figure 10 for 6 cells having various degree of EIsf. The maps are arranged in the descending order of EIsf (indicated by a number at the upper-right corner of each map). For these 6 cells, various extent of elongation was observed from a small value of EIsf (1.10) to a large value (3.66). It suggests that there are multiple grades in integrating process of SF channels from one cell to another. Results for the previous simple cell (Fig. 8E-H) are shown in Figure 10F. This cell shows almost a circular shape of the binocular interaction map, suggesting the lack of integration.

EIsf from all neurons in my sample are summarized in Figure 11A-C. To examine a possible relationship between the degree of elongation and the cell type (simple/complex), EIsf and F1/F0 ratio measured with drifting gratings for each neuron are plotted (Fig. 11A). Blue dots indicate simple cells while red dots represent complex cells. Statistical significance of elongation was tested for all the data, and those that are significant are indicated in deeper color (EIsf > 1.0, bootstrap test, $p < 0.01$). A strong negative correlation between these two values was observed ($r = -0.55$, $p < 0.001$), suggesting that integration of SF channels is essentially absent for simple cells. Complex cells generally had large degree of elongation in comparison, but there were also some complex cells without elongation or small degree of elongation.

Notice that even among simple cells, there are some with significant elongation. I wondered if those cells might have complex-like mechanisms internally despite being classified as simple based on F1/F0 ratio. To examine this discrepancy in more detail, I also examined binocular receptive field (bRF) structure for each neuron. A separability index of a bRF is another metric that reflects the "simple-ness" based on binocular data, and appears to be a better metric as an indication of subunits (Sanada and Ohzawa, 2006). In contrast, F1/F0 ratio is usually defined monocularly. A separability index of a bRF being 1.0 indicates a totally separable map while 0 indicates an inseparable map. An ideal simple cell with a linear Gabor-like RF for each eye will have the index of 1.0. The detailed procedure for obtaining a bRF of each cell is explained later. Figure 11B presents the relationship between the separability index of a bRF and EIsf, showing significant negative correlation ($r = -0.61$, $p < 0.01$). For these results, simple cells with EIsf > 1.0 tended to show rather complex-like (inseparable) bRFs although they all had modulated responses to drifting grating stimuli. Such neurons may have intermediate property between simple and complex cells, as reported in the past studies (Sanada and Ohzawa, 2006; Sasaki et al., 2010).

Since one of the expected consequences of integration of SF channels for disparity detection is a widening of SF bandwidth, correlation between the SF bandwidth and EIsf would be predicted. Therefore, I examined a relationship between the SF bandwidth measured with drifting grating stimuli and EIsf. There appears to be a correlation between these two values, although it was not significant ($r = 0.23$, $p =$

0.057; Fig. 11C).

The previous study that defined the separability index of a bRF as noted above (Sanada and Ohzawa, 2006), primarily examined a difference of optimal SFs between the left and right eyes, possibly encoding three-dimensional slant. Is there any relationship between elongation in the SF domain and difference in the left and right optimal SF? The binocular SF interaction map of such a neuron may be elongated off the exact diagonal. First, I examined this. Figure 11D shows the distribution of left and right optimal SF difference for four groups of neurons. Upper row in Figure 11D represents cells with significant elongation in the SF domain (EIsf > 1.0) while the lower histograms show data without elongation. The Data confirm the presence of cells with significant left-right optimal SF difference, indicated by deeper colors in the histograms. Among 50 neurons (10 simple, 37 complex, 3 unclassified) with significant elongation, 18 neurons showed significant SF difference between the two eyes. No particular tendency or relationship was found among the degree of elongation, cell type (simple/complex), and offset of binocular SF interaction maps from the diagonal (not shown).

Note that even if there is a difference in the optimal SFs between the two eyes, elongated binocular SF profiles should satisfy a specific relationship for maintaining the ability to signal consistent surface slant in depth (Sanada and Ohzawa, 2006). Specifically, the left-right SF ratio should remain constant across different binocular SF pairs. Next, this was examined as follows for eleven of neurons showing significant left-right SF difference with EIsf >1.5. I fitted binocular SF interaction maps of such neurons with an elongated 2-dimensional Gaussian function, allowing its long axis to tilt from 45° diagonal. After fitting, two high and low left-right SF combinations along the fitted long axis were selected, which was separated by a distance of 0.8 sigma from the fitted center. Left-right SF ratio was then calculated at each point and compared. Although the number of data was small (n=11), there was a significant correlation of left-right SF ratios between low and high SF combinations, suggesting that these neurons might integrate signals while preserving consistency of surface slant information (Pearson's correlation coefficient, $r = 0.77$, $p < 0.05$).

In Figure 12, the relationships among the other basic parameters are shown. Figure

12A and B present the relation of optimal SF measured with monocular stimuli to EIsf and disparity frequency, respectively. There was a weak positive correlation between the optimal SF and elongation in the SF domain, suggesting that neurons tuned to higher SF may carry more reliable disparity signal than neurons tuned to lower SF ($r = 0.24$, $p < 0.05$). Furthermore, the scatter plot in Figure 12B shows a tendency also shown in previous studies (Ohzawa et al., 1997; Prince et al., 2002; Read and Cumming, 2003) where there was a strong correlation, but the optimal SF generally is higher than optimal disparity frequency ($r = 0.73$, $p < 0.001$). Figure 12C and D show the relationships of optimal orientation and its bandwidth to the EIsf. No special relationships were observed between the elongation in the SF domain and the optimal orientation or its bandwidth.

Reconstruction of binocular receptive fields (bRFs) and disparity tunings in the space domain

So far, I have shown that substantial elongation of binocular SF profiles occurs in many A17 disparity-tuned neurons. However, the most important point for accurate disparity detection is selectivity to constant inter-ocular position shift across all the SF components as described in Introduction. In other words, a neuron which has sharp and a diagonally elongated binocular profile in the SF domain, must share the same preferred disparity across different SF bands (Wagner and Frost 1993, 1994). To address this question, the SF domain analysis is not appropriate, and we must return to the space domain where binocular disparity tuning may be obtained directly.

A disparity tuning curve was calculated in two steps as follows. First, a reverse correlation analysis in the joint left-right space domain was performed to obtain a binocular receptive field (bRF) using methods similar to previous studies (Anzai et al., 1999; Sasaki et al., 2010). Spike-triggered grating pairs were selected for an optimal correlation delay (Fig. 13A), and one-dimensional spike-triggered sin-waves were multiplied between the two eyes to produce binocular interaction terms (Fig. 13B). If contrasts are the same polarity between the two eyes (white-white or black-black), value of the interaction term becomes positive. On the other hand, when contrasts between the two eyes are opposite (white-black or black-white), the binocular interaction term

becomes negative. By repeating the above calculation of binocular interaction terms for all spikes and summing them, a bRF was obtained in the joint left-right space domain (Fig. 13C). In this bRF map, binocular disparity is constant along the +45° diagonal while it changes along −45° diagonal. Therefore, in the second step, a disparity tuning curve is computed by integrating the map along the +45° diagonal.

Figure 13D–I shows the results of the first step, i.e., bRFs for 6 example neurons because they are inherently much more informative than reduced one-dimensional disparity tuning curves. In each map, red color indicates binocular response to the same contrast polarity for the two eyes, while blue represents that for the opposite contrast polarity. Horizontal and vertical axes define the position along the axis orthogonal to preferred orientation in the left and right eye, respectively.

Figure 13D shows a typical bRF for a simple cell. Because of the selectivity to monocular phase, the map shows a separable profile. On the other hand, bRFs for other five complex cells show inseparable shape oriented along the constant disparity line (Fig. 13E–I). The bRF in Figure 13E shows an inverted bRF profile, found relatively rarely, for which the strongest subregion is to a combination of opposite polarity contrasts across the eyes. If the central blue region is at zero disparity, it would be a tuned-inhibitory cell. Because my preparation was anesthetized and paralyzed, I do not have accurate information on retinal correspondence. Figure 13F and G show even-symmetric and odd-symmetric bRFs, respectively. The maps in Figure 13H and I represent bRFs for the neuron that showed large degree of elongation in the SF domain with EIsf 3.42 and 3.66 for H and I, respectively. Notice that, for these neurons, the reconstructed bRFs in the space domain (H and I) are also thin and highly elongated along +45° diagonal. Analysis on the relationship between the SF domain and the space domain for elongation is described later.

Are optimal disparities matched across different SF bands?

Having described the 2-dimensional bRFs, I now return to the original question and examine binocular disparity tuning curves. Are combined elements tuned to the same common disparity across different SF sub-bands? For this purpose, I note that the bRF and disparity tuning curve may be calculated for a subset of stimuli limited to a

particular SF band. This is achieved by simply limiting the spike-triggered stimuli to within an arbitrary SF band during a reverse correlation analysis for constructing bRFs.

To compare disparity selectivity between mechanisms tuned to high and low SF, bRFs for two SF bands were obtained by using only upper half or lower half of the SF components. Figure 14 shows bRFs and disparity tuning curves for a representative complex cell (the same cell shown in Fig. 8A–D and Fig. 10B). SF ranges used for each reconstruction are illustrated in the upper row of Figure 14A–C in the form of binocular SF interaction maps. SF components outside the selected SF band are indicated by dark blue, and each bRFs are shown under the respective SF interaction map (A: all SF components were used, B: only lower half of SFs were used, and C: only upper half of SFs were used).

In Figure 14D, computed disparity tuning curves of each bRFs are superimposed (gray: all SF components, red: low SF components, and blue: high SF components). Clearly, peaks of all the tuning curves are at the same disparity. The result suggests that this neuron combines inputs across different SFs while maintaining exact tuning to the same disparity. Such a processing may allow robust detection of disparity, because responses to the false matches at side-lobes of a tuning curve may be reduced due to mutual cancellation of side-lobes of tuning curves for different SF bands (Fig.14D). However, notice that the care is needed for this analysis about interpretation of side-lobes, because based on an uncertainty principle (Daugman, 1985), bandwidth limitation in the SF domain generates increased number of bounces in side-lobes in the space domain. The bandwidth-limited reconstructions of bRFs were carried out using approximately 1.5 octave bandwidth in Figure 14B and C (a half of 3 octaves total). Although this is close to the average bandwidth of A17 neurons, at least part of side-lobes might have been due to the bandwidth limitation.

If an integration of disparity detectors in the SF domain always occurs under the constraint of common disparity for all SFs, neurons with refined binocular SF profiles should show similar results to that shown in Figure 14. Therefore, I next examined whether this constraint is common in other neurons that showed large elongation in the SF domain. Neurons with EIsf larger than 1.5 were examined with the same analysis as that for Figure14, and the peak difference of disparity tuning curves between low and high SF bands was evaluated. Figure 15A illustrates the calculation of ‘normalized peak

difference,' that indicates the quantitative index of the difference. ΔPeak is the separation between the peaks of tunings between the different SF bands, and normalized peak difference is defined as the ΔPeak divided by the wavelength (1/SF) of disparity tuning curve (all SF condition). Figure 15B shows the distribution of normalized peak difference and its relationship with the phase of original disparity tuning curve for my sample of cells (n = 44). As the histogram at the top of Figure 15B shows, most cells had a relatively small peak difference of less than 0.1 between the low and high SF bands, indicating a good general match of encoded disparity across different SF sub-bands. Therefore, for the majority of disparity coding neurons, combined binocular SF pairs satisfy the rule of sharing the same preferred disparity. However, for a small fraction of the population, this was not the case. I wondered here if the alignment of peak disparities might depend on the symmetry of disparity tuning curves. This is because the peak alignment is the same as centering alignment (of the envelopes) for even-symmetric disparity tuning curves as illustrated in Figure 15A (blue and red curves). On the other hand, in order to peak-align odd-symmetric disparity tuning curves across different SF bands, centers of disparity tuning curves must be offset accordingly. In this case, it is conceivable that some neurons may combine inputs across SF sub-bands with "zero-crossing" alignment presumably for different purposes. The result for an example neuron that shows such an alignment across different SF bands is indicated in Figure 15C. I do not have a definite conclusion on this, because there are only small numbers of neurons with a large normalized peak difference (> 0.1). However, those all had nearly odd-symmetric disparity tuning curves as indicated by the phase of disparity tuning curves close to 90° or 270° (Fig. 15B).

Relationship between pooling in the SF domain and the space domain

Recall that some of the reconstructed bRFs showed highly elongated profiles in the space domain (Fig. 13). As previous studies show, some degree of spatial pooling of multiple disparity detectors occurs for a portion of A17 neurons (Sasaki et al. 2010). However, the degree of spatial pooling found in previous studies (Sasaki & Ohzawa, 2007; Sasaki et al., 2010) is relatively limited compared with the elongation found for some neurons in this study (Fig. 13). Therefore, I have examined a possibility that the pooling, as evaluated by elongation of binocular SF interaction maps, is not completely

independent in the space and the SF domains. Specifically, I consider if the apparent elongation of bRFs in space may arise from pooling in the SF domain. If pooling across multiple SF bands can reduce sidebands in disparity tuning curves, it may essentially restrict the extent of bRFs along the –45° axis without necessarily extending bRFs along the +45° axis (X axis). To examine this possibility, I consider a disparity detector that pools only in the SF domain. Such a detector may be constructed by summing the output of multiple disparity energy units aligned at the center (Fig. 16). The bRF of the pooled units is slightly elongated along the constant disparity (+45°) axis without any pooling in the space domain (gray elongated contour in Fig. 16B). Therefore, pooling purely in the SF domain itself can be responsible for a part of extensive spatial pooling along the X axis.

However, bRFs constructed by such SF pooling alone shows some distortion of the bRF not observed in the bRFs of actual neurons, as shown in Figure 16B. Because the bRF of a unit tuned to lower SF covers a larger space than that tuned to higher SF (Fig. 16A), the pooled bRF becomes sharpened only at the center, but remaining broad at the edges shown by superimposed tuning curves of cross-sections at the center and the edge in Figure 16B (bottom). On the other hand, actual neurons showed no such distortion of bRFs (Fig. 16C). Tuning curves at the center and the edge are almost identical for these neurons, although they show highly elongated shape of bRFs. Therefore, some degree of spatial pooling clearly occurs for neurons shown in Figure 16C and others in our sample. A prediction from this analysis is that neurons with broad-band SF tuning tend to have a corresponding degree of spatial pooling for maintaining consistent disparity tuning across all locations of bRFs.

Likewise, I must also consider a possibility in the opposite direction where the apparent elongation of binocular SF interaction map may be caused by pooling multiple disparity energy units in the space domain. In Figure 17, a schematic explanation is provided for such elongation in the SF domain. Consider a neuron that pools multiple disparity energy units at different spatial positions along the X axis (broken gray circles), but tuned to the same SF (light red circle) and to common disparity. When SFs are identical between the two eyes (45° broken line), inter-ocular phase difference is always the same along the X axis. However, if left and right SFs are different, inter-ocular phase difference changes depending on the positions along the X axis as shown at the

bottom right of Figure 17. Such differences produce different firing rates across energy units at different positions, which means that these energy units do not fire in a concerted manner. Thus, the strength of binocular SF interaction is reduced most along the –45° axis where SF is different maximally between the two eyes, generating the apparent elongation of binocular SF interaction maps (elongated red filled region along +45° diagonal in Fig. 17).

Figure 18 illustrates the relationship between pooling in the SF domain and that in the space domain. I computed the elongation index along the X axis (EIx) for each bRF to capture the degree of pooling in the joint left-right space domain. For EIx, the aspect ratio of the Fourier spectrum was calculated for each bRF (Sanada et al., 2006). There was a strong positive correlation between these two values ($r = 0.82$, $p < 0.001$), probably due to complex interactions between the SF domain and the space domain described above. Pooling in one domain generates a degree of apparent elongation of the joint profile in the other domain, making the measurement difficult or almost impossible for estimating the degree of effect caused in a single domain independently. The profile I measured here may be the end product generated by mixed pooling in both SF and position domains including complex interactions.

Binocular matching with substantially sharp SF tunings

Although the elongation of binocular SF interaction profiles appears to be caused by complex factors including pooling in both SF and space domains, consequences of the elongation for the detection of binocular disparity may be considered independently. Since the analysis presented in Figure 17 suggests that pooling in either domain may cause the narrowing of SF interaction profiles, I have examined the SF bandwidths in both monocular and binocular conditions. Specifically, I ask whether binocular SF bandwidth is narrower than monocular SF bandwidth. To address this question, a cross-section of a binocular SF interaction map for the dominant eye (Fig. 19A, B) was compared with monocular SF tuning curve obtained using drifting grating stimuli (broken black curve in Fig. 19B). To clarify, for example, horizontal cross-sections of the profile in Figure 19A represent SF tuning curves in the left eye in the presence of constant various SFs in the right eye. These cross-sections are plotted in Figure 19B as

colored solid curves. Population data show that many neurons have binocular SF bandwidth narrower than 0.5 octave width (Fig. 19C). The result suggests that binocular matching is conducted with much higher SF precision than that predicted from average monocular SF bandwidth of neurons in A17.

DISCUSSION

In this study, I examined how the integration of information from different SF bands is achieved in disparity-sensitive binocular neurons in the striate cortex. A previous computational study proposed a model of combining outputs from multiple energy units at different scales for robust estimation of binocular disparity (Fleet et al., 1996). Physiologically, a previous study indicated that multiple excitatory and suppressive subunits contributed to generate disparity-selective responses in neurons of monkey V1 (Tanabe et al., 2011). However, no direct assessment has been conducted physiologically to evaluate the integration of multiple excitatory channels tuned to different SFs as illustrated in Figure 6. I obtained a binocular SF interaction map, which reflects the degree of integration by analyzing phase-based disparity tunings in the joint left-right SF domain. A subset of complex cells showed substantial elongation of binocular SF interaction profile along the 45° diagonal, which is not predicted from a single energy unit.

Interactions of pooling across multiple stimulus dimensions

As noted in the last part of Results, a complex interaction between SF and space domains occurs for the binocular profile when pooling of multiple energy units are considered. In theory, other forms of interactions are conceivable across multiple stimulus dimensions. For example, given that neurons in the primary visual cortex are tuned sharply for orientation, pooling in the orientation dimension might also affect some of the properties we examined in the present study. For the complete understanding of underlying neural mechanisms, it may be desirable to consider and measure the degree of pooling by taking all of such relevant domains into account exhaustively, although practically it does not seem so easy. Comprehensive and systematic examination for such complicated interactions may be needed in the future study, both theoretically and physiologically.

Binocular matching is conducted with substantially sharp SF tunings

The results presented in Figure 19 suggest that binocular matching is conducted with

much higher SF precision than that predicted from average monocular SF bandwidth of neurons in A17. What benefits are there for narrow-band binocular matching? Receptive fields of typical A17 neurons appear to be designed for achieving a best balance between high positional (x, y) and SF (fx, fy) accuracies, giving typical monocular SF bandwidth of 1.3 octaves. However, when matching left and right patterns, some neurons in my data show that the binocularly relevant SF bandwidth is around 0.5 octaves. It means that even though the front-end linear filter, simple cells, and complex cells in the next stage are of medium bandwidth monocularly, a subset of the same complex cells achieves narrow-band binocular matching. It is somewhat surprising that these widely different bandwidths can coexist in the same cells.

One of the benefits of narrow-band binocular matching may be that, in the SF domain, signaling accuracy with respect to spatial phase is particularly important, as that is the parameter which is directly related to positional shift. However, such a conversion between phase and positional shift relies on the accurate knowledge of SF. Narrow-band SF tuning that applies only for binocular matching appear to achieve the desired condition, without necessarily modifying basic front-end tuning characteristics of neurons. Obviously, the output of these initial stage neurons are used for many purposes other than stereopsis, such as determining shape and motion velocities. For those other visual functions, extremely narrow-band SF tuning may not be desirable. Overall, the binocular processing for stereopsis appears to be remarkably well-designed with highly suitable tuning properties actually achieved for different purposes, but with relatively simple schemes.

Relation to coarse-to-fine mechanism

Some of computational models of stereoscopic processing take a sequential approach, known as ***coarse-to-fine*** algorithms (Marr and Poggio, 1979; Quam, 1987; Chen and Qian, 2004; Li and Qian, 2015). In these algorithms, disparity information is hierarchically processed from coarse to fine scales, improving the accuracy of disparity detection as it proceeds. Do cortical neurons also implement a sequential refinement in integrating multiple SF bands? While I cannot directly address this question, my analysis implies that neurons simply pool the output of multiple subunits tuned to different SFs but with common preferred disparity. No explicit and non-trivial

sequential interactions from coarse-to-fine (or from low SF to high SF) are assumed in this pooling scheme. Such a simple pooling mechanism appears sufficient to explain the results obtained.

However, there may be a sequential element of coarse-to-fine organization in the simple pooling process. It is well known that, in monocular spectral-time receptive field analyses, the optimal SF increases from low to high as a function of response delay (Bredfeldt and Ringach, 2002; Mazer et al., 2002; Nishimoto et al., 2005). In other words, signals for high SFs arrive at the neuron with a longer temporal delay than low SF signals. The same phenomenon is also observed for binocular responses. Disparity-time response analyses reveal progressive shift of disparity frequency as a function of temporal delay (Menz and Freeman, 2003). My data also showed similar tendency in the time course of disparity tuning (not shown). Specifically, SF interaction map showed elongation in a single time delay (e.g., Fig. 9A), and at the same time, showed slight progressive shift of its optimal SF from low to high as the temporal delay was increased. Therefore, regardless of the implementation, the real visual system may also effectively achieve progressive computation and refinement of disparity information simply by summing signals from multiple subunits with various temporal delays.

Possible effects of suppressive elements

Although I have so far assumed that the elongated shape of a binocular SF interaction map is caused by pooling of multiple excitatory channels, possibility of suppression effect may not be negligible (Tanabe et al., 2011). Instead of elongating a response region by adding subunits (Fig. 6B), it is also possible that binocular response is inhibited at specific SF combinations where difference of SF is large between the two eyes. Unfortunately, because my measurements contain too few blank stimuli to estimate the baseline response level accurately, it is difficult to evaluate suppressive responses for this purpose. However, although possible in principle, such a hypothetical scheme seems unlikely or at least inefficient, because units exerting suppression need to be constructed in the first place with inputs from highly unmatched SFs between the two eyes. Such neurons with a large difference in preferred SFs are not generally found (Sanada and Ohzawa, 2006).

Other schemes are also possible if one allows more complexity in the model. For example, inhibitory input from a neuron tuned to a slightly higher SF than the excitatory neuron may be present. If such an inhibitory neuron is turned off (assuming an appropriate disparity tuning), the result will be a disinhibition, which will result in net increase in excitation for the postsynaptic neuron. However, such schemes are indistinguishable from excitatory input as recorded from the postsynaptic neuron. Therefore, functionally, such schemes fall within the framework being considered. Unless there is a substantial functional difference, it is simpler and more natural to consider a detector that combines multiple excitatory subunits rather than one that use suppression.

Alignment of disparity tuning curves for pooling across SF bands

Majority of neurons in my sample exhibit elongation of binocular profiles in the SF domain while preserving a common optimal disparity across SF bands. As a previous computational study shows (Fleet et al., 1996), combining energy units tuned to different SFs but to common disparity would improve detection accuracy by increasing response probability at the true disparity and decreasing it at the false disparities. However, there are some neurons that showed a large difference in optimal disparity when pooling across SF bands (Fig. 15B). These neurons tended to have odd-symmetric disparity tuning curves, and the results of multi-SF-band analysis show that disparity tuning curves for high and low SFs are aligned approximately at zero-crossings rather than at peaks (Fig. 15C). Do such neurons play some functional roles?

In higher visual areas, especially those in the dorsal visual stream such as MT and MST, it is known that the number of disparity-selective neurons with odd-symmetric tuning curve substantially increases (Cumming and DeAngelis, 2001; DeAngelis and Uka, 2003). A possible role is suggested that the output of these neurons provides a signal for oculomotor vergence control (Masson et al., 1997). Characteristics of signals required for fine vergence control may be different from those for depth perception. Specifically, it may be more important to achieve high sensitivity near 0 disparity for determining the direction and size of vergence: converge or diverge, rather than detecting a disparity. Neurons with odd-symmetric disparity can provide such a signal, although the same information may be obtained from peaks of multiple neurons. Such a

vergence signal should also function for large disparity deviations. In other words, SF bands may be pooled such that the slope at zero-crossing is increased and the output maintained even at large disparities. This goal is achieved by pooling multiple SF components with zero-crossing alignment, as the famous Fourier series decomposition of a square wave indicates.

Effects of SF pooling on the shape of disparity tuning curve

There are two known refinements in the shape of disparity tuning curves in neurons beyond that of a disparity energy model. The refinement begins in A17 (Ohzawa, 1998; Haefner and Cumming, 2008; Tanabe et al., 2011) but become more pronounced in high-order areas such as V4 and IT. One is the reduction of multiple side-lobes as illustrated in Figure 14, and noted by Fleet et al. (1996). The other is the reduction of responses to anti-correlated random-dot stereograms (aRDS) (Janssen et al., 2003; Tanabe et al., 2004). These two factors are often discussed together but are distinct (Nieder and Wagner, 2001). In relation to the latter, a previous study shows that in monkey V4, disparity-sensitive neurons show a correlation between SF bandwidth and the degree of attenuation of response amplitude for aRDS compared with that for correlated RDS (cRDS) (Kumano et al., 2008). They interpret this as a consequence of integrating multiple SF channels. However, pooling multiple disparity-selective units across SF bands by itself doesn't produce attenuation of response amplitude to aRDS. Responses to aRDS are merely inverted versions of disparity tuning for cRDS.

REFERENCES

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

Barlow HB, Blakemore C, Pettigrew JD (1967) The neural mechanism of binocular depth discrimination. J Physiol (Lond) 193:327–342.

Bredfeldt CE, Ringach DL (2002) Dynamics of spatial frequency tuning in macaque A17. J Neurosci 22:1976–1984.

Chen Y, Qian N (2004) A coarse-to-fine disparity energy model with both phase-shift and position-shift receptive field mechanisms. Neural Comput 16:1545–1577.

Cumming BG, DeAngelis GC (2001) The physiology of stereopsis. Annu Rev Neurosci 24:203–238.

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

DeAngelis GC, Ohzawa I, Freeman RD (1993b) 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.

DeAngelis GC, Uka T (2003) Coding of horizontal disparity and velocity by MT neurons in the alert macaque. J Neurophysiol 89:1094–1111.

Doi T, Fujita I (2014) Cross-matching: a modified cross-correlation underlying threshold energy model and match-based depth perception. Front Comput Neurosci 8:127.

Ferster D (1981) A comparison of binocular depth mechanisms in areas 17 and 18 of the

cat visual cortex. J Physiol (Lond) 311:623–655.

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

Freeman J, Simoncelli EP (2011) Metamers of the ventral stream. Nat Neurosci 14:1195–1201.

Freeman RD, Robson JG (1982) A new approach to the study of binocular interaction in visual cortex: normal and monocularly deprived cats. Exp Brain Res 48:296–300.

Gattass R, Sousa APB, Rosa MGP (1987) Visual topography of V1 in the Cebus monkey. J Comp Neurol 259:529–548.

Gross CG, Rocha-Miranda CE, Bender DB (1972) Visual properties of neurons in inferotemporal cortex of the Macaque. J Neurophysiol 35:96–111.

Haefner RM, Cumming BG (2008) Adaptation to natural binocular disparities in primate A17 explained by a generalized energy model. Neuron 57:147–158.

Helmholtz H (1910) Handbuch der physiologischen Optik 3rd edition (A Gullstrand, J von Kries, W Nagel, eds), Leopold Voss, Hamburg.

Hubel DH, Wiesel TN (1959) Receptive fields of single neurones in the cat's striate cortex. J Physiol 148:574–591.

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

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

Kumano H, Tanabe S, Fujita I (2008) Spatial frequency integration for binocular correspondence in macaque area V4. J Neurophysiol 99:402–408.

Lee TS (1996) Image Representation Using 2D Gabor Wavelets. IEEE Trans Pattern Anal Mach Intell 18:959–971.

Li Z, Qian N (2015) Solving stereo transparency with an extended coarse-to-fine disparity energy model. Neural Comput in press.

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

Masson GS, Busettini C, Miles FA (1997) Vergence eye movements in response to binocular disparity without depth perception. Nature 389:283–286.

Maunsell JHR, Newsome WT (1987) Visual Processing in Monkey Extrastriate Cortex. Annual Review of Neuroscience 10:363–401.

Maunsell JH, Van Essen DC (1983) Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. J Neurophysiol 49:1148–1167.

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

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

Movshon JA, Thompson ID, Tolhurst DJ (1978a) Spatial summation in the receptive fields of simple cells in the cat's striate cortex. J Physiol (Lond) 283:53–77.

Movshon JA, Thompson ID, Tolhurst DJ (1978b) Receptive field organization of complex cells in the cat's striate cortex. J Physiol 283:79–99.

Movshon JA, Thompson ID, Tolhurst DJ (1978c) Spatial and temporal contrast sensitivity of neurones in areas 17 and 18 of the cat's visual cortex. J Physiol (Lond) 283:101–120.

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

Ninomiya T, Sanada TM, Ohzawa I (2012) Contributions of excitation and suppression in shaping spatial frequency selectivity of V1 neurons as revealed by binocular measurements. J Neurophysiol 107:2220–2231.

Nishimoto S, Arai M, Ohzawa I (2005) Accuracy of subspace mapping of spatiotemporal frequency domain visual receptive fields. J Neurophysiol 93:3524–3536.

Nishimoto S, Ishida T, Ohzawa I (2006) Receptive field properties of neurons in the early visual cortex revealed by local spectral reverse correlation. J Neurosci 26:3269–3280.

Nishio A, Goda N, Komatsu H (2012) Neural selectivity and representation of gloss in the monkey inferior temporal cortex. J Neurosci 32:10780–10793.

Ohzawa I (1998) Mechanisms of stereoscopic vision: the disparity energy model. Curr Opin Neurobiol 8:509–515.

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

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

Ohzawa I, Freeman RD (1986) The binocular organization of complex cells in the cat's visual cortex. J Neurophysiol 56:243–259.

Orban GA (1997) Visual Processing in Macaque Area MT/V5 and Its Satellites (MSTd and MSTv). In: Extrastriate Cortex in Primates (Rockland KS, Kaas JH, Peters A, eds), pp 359–434 Cerebral Cortex. Springer US.

Poggio GF, Fischer B (1977) Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. J Neurophysiol 40:1392–1405.

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

Qian N (1994) Computing Stereo Disparity and Motion with Known Binocular Cell Properties. Neural Comput 6:390–404.

Quam LH, (1987) Hierarchical warp stereo. In: Readings in Computer Vision (MA Fischler, O Firschein, eds), pp 80-86 Morgan Kaufmann Publishers Inc. San Francisco, CA, USA.

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

Ringach DL, Sapiro G, Shapley R (1997) A subspace reverse-correlation technique for the study of visual neurons. Vision Res 37:2455–2464.

Sanada TM, Ohzawa I (2006) Encoding of three-dimensional surface slant in cat visual areas 17 and 18. J Neurophysiol 95:2768–2786.

Sasaki KS, Ohzawa I (2007) Internal spatial organization of receptive fields of complex cells in the early visual cortex. J Neurophysiol 98:1194–1212.

Sasaki KS, Tabuchi Y, Ohzawa I (2010) Complex cells in the cat striate cortex have multiple disparity detectors in the three-dimensional binocular receptive fields. J Neurosci 30:13826–13837.

Scharstein D, Hirschmüller H, Kitajima Y, Krathwohl G, Nešić N, Wang X, Westling P (2014) High-Resolution Stereo Datasets with Subpixel-Accurate Ground Truth. In: Pattern Recognition (Jiang X, Hornegger J, Koch R, eds), pp 31–42 Lecture Notes in Computer Science. Springer International Publishing.

Schor CM, Wood I (1983) Disparity range for local stereopsis as a function of luminance spatial frequency. Vision Res 23:1649–1654.

Simoncelli EP, Heeger DJ (1998) A model of neuronal responses in visual area MT. Vision Research 38:743–761.

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

Tanabe S, Haefner RM, Cumming BG (2011) Suppressive mechanisms in monkey V1 help to solve the stereo correspondence problem. J Neurosci 31:8295–8305.

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

Tanaka K (1996) Inferotemporal cortex and object vision. Annu Rev Neurosci 19:109–139.

Tsao DY, Freiwald WA, Tootell RBH, Livingstone MS (2006) A Cortical Region Consisting Entirely of Face-Selective Cells. Science 311:670–674.

Uka T, Tanaka H, Yoshiyama K, Kato M, Fujita I (2000) Disparity selectivity of neurons in monkey inferior temporal cortex. J Neurophysiol 84:120–132.

Valois RLD, Albrecht DG, Thorell LG (1978) Cortical Cells: Bar and Edge Detectors, or Spatial Frequency Filters? In: Frontiers in Visual Science (Cool DSJ, III DELS, eds), pp 544–556 Springer Series in Optical Sciences. Springer Berlin Heidelberg.

Wagner H, Frost B (1994) Binocular responses of neurons in the barn owl's visual Wulst. J Comp Physiol A 174:661–670.

Wagner H, Frost B (1993) Disparity-sensitive cells in the owl have a characteristic disparity. Nature 364:796–798.

Westheimer G, McKee SP (1980) Stereogram design for testing local stereopsis. Invest Ophthalmol Vis Sci 19:802–809.

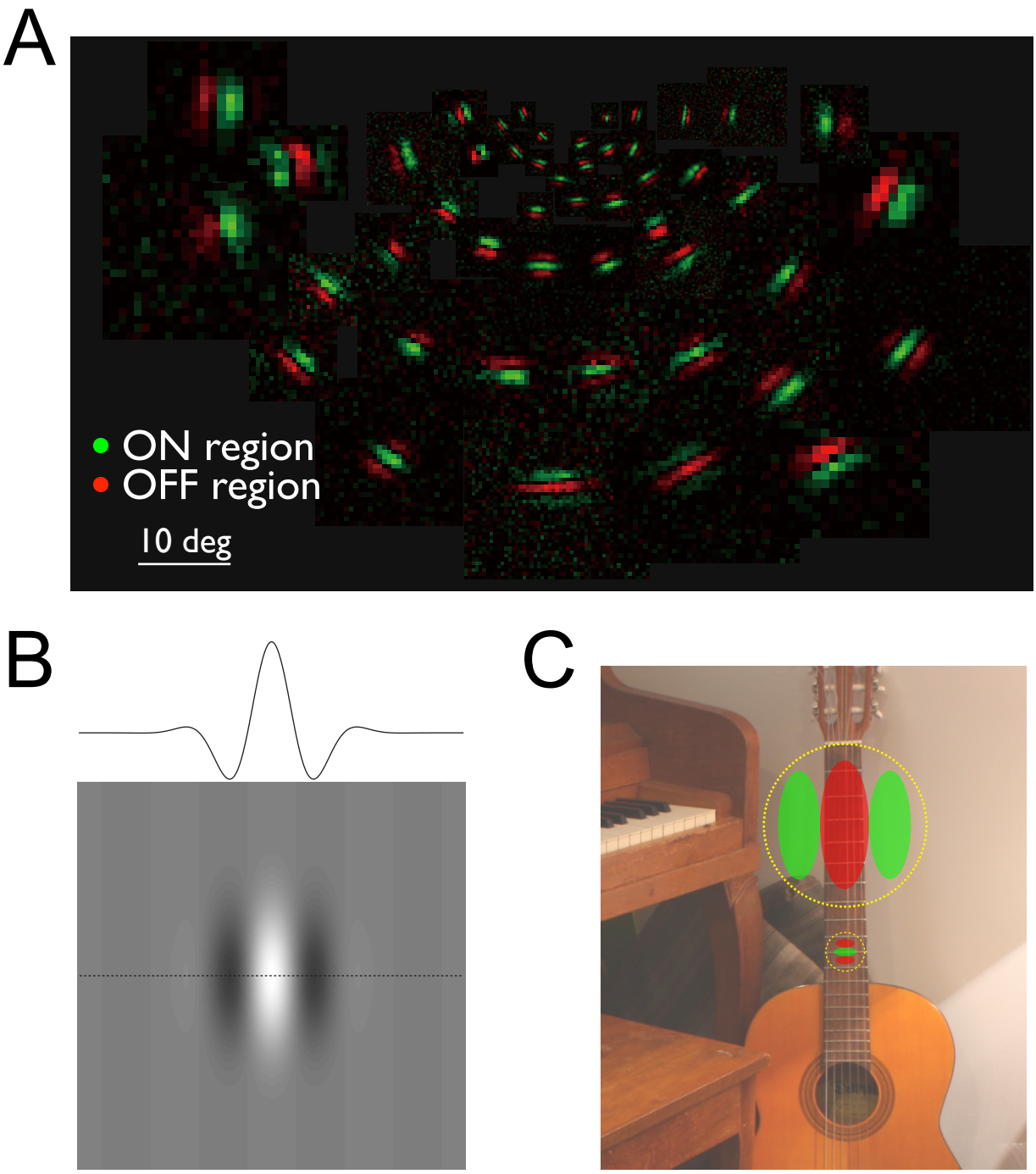

Figure 1. ***A***, A variety of receptive fields (RFs) of simple cells measured in cat A17 are shown. Green-colored area indicates an ON region and red-colored area indicates an OFF region in each RF. RFs are arranged according to their optimal orientation and SF (This panel is provided by Sasaki. KS). As a neural population, A17 simple cells construct a self-similar wavelet pyramid. ***B***, Even-symmetric Gabor function is represented in the two-dimensional form in the grayscale image. Its horizontal cross-section that passes through the center is indicated at the top. The RF structure of A17 simple cell is well fitted by such a function. ***C***, A schematic illustration of visual feature detection by Gabor wavelets is shown. Two of the Gabor-shape RFs are superimposed on the different feature at different parts of the image (surrounded by broken yellow circles). A horizontal fine feature (e.g., frets of a guitar) is detected by an RF with horizontal orientation and high SF, while a vertical coarse feature (e.g., a neck of a guitar) is detected by an RF with vertical orientation and low SF.

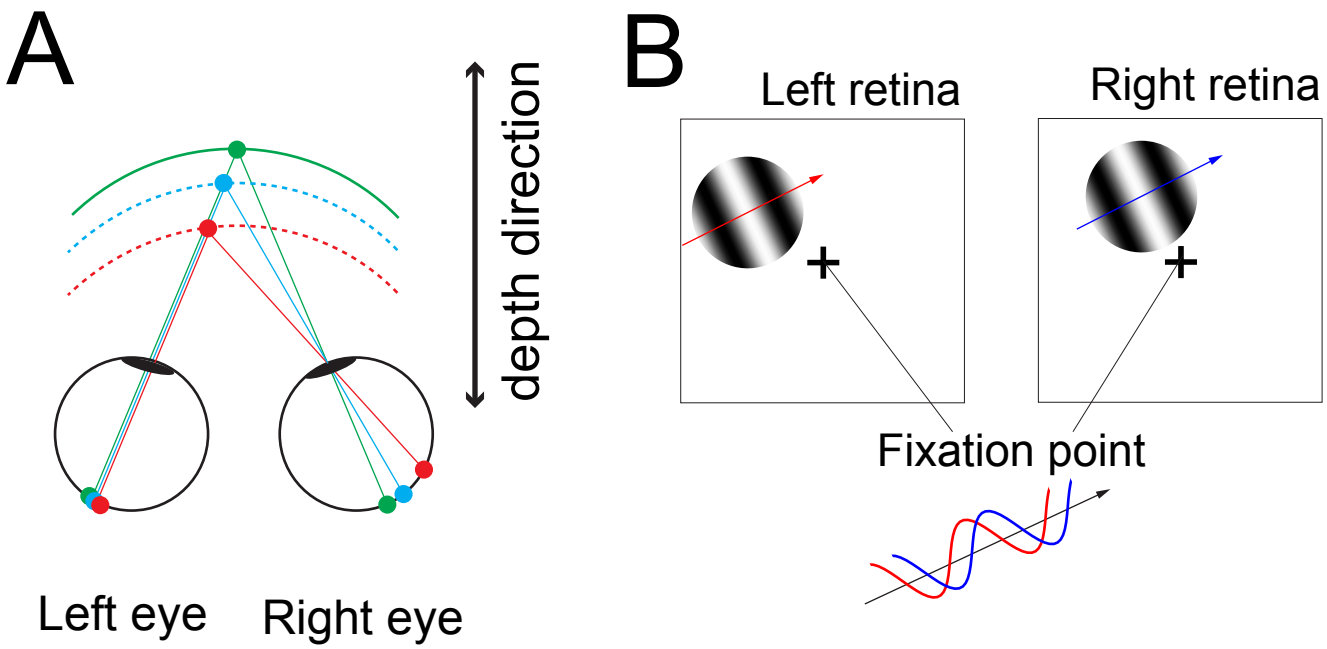

Figure 2. ***A***, Geometrical relationship between three-dimensional depth and the retinal position shift of a projected object is illustrated. Three different locations along depth direction are indicated in different colors. As location of an object is shifted along depth direction, a relative position and resulting left-right difference between the two retinal projections change systematically. ***B***, Direction of binocular disparity produced by oriented two-dimensional sine-waves is indicated. Top panel shows grating stimuli projected onto portions of the retinae (illustrated as squares). Positional shift of sine-wave patterns between the left and right retinae is detectable only along the direction orthogonal to its orientation (shown at the bottom), since there is no change of luminance along its orientation direction.

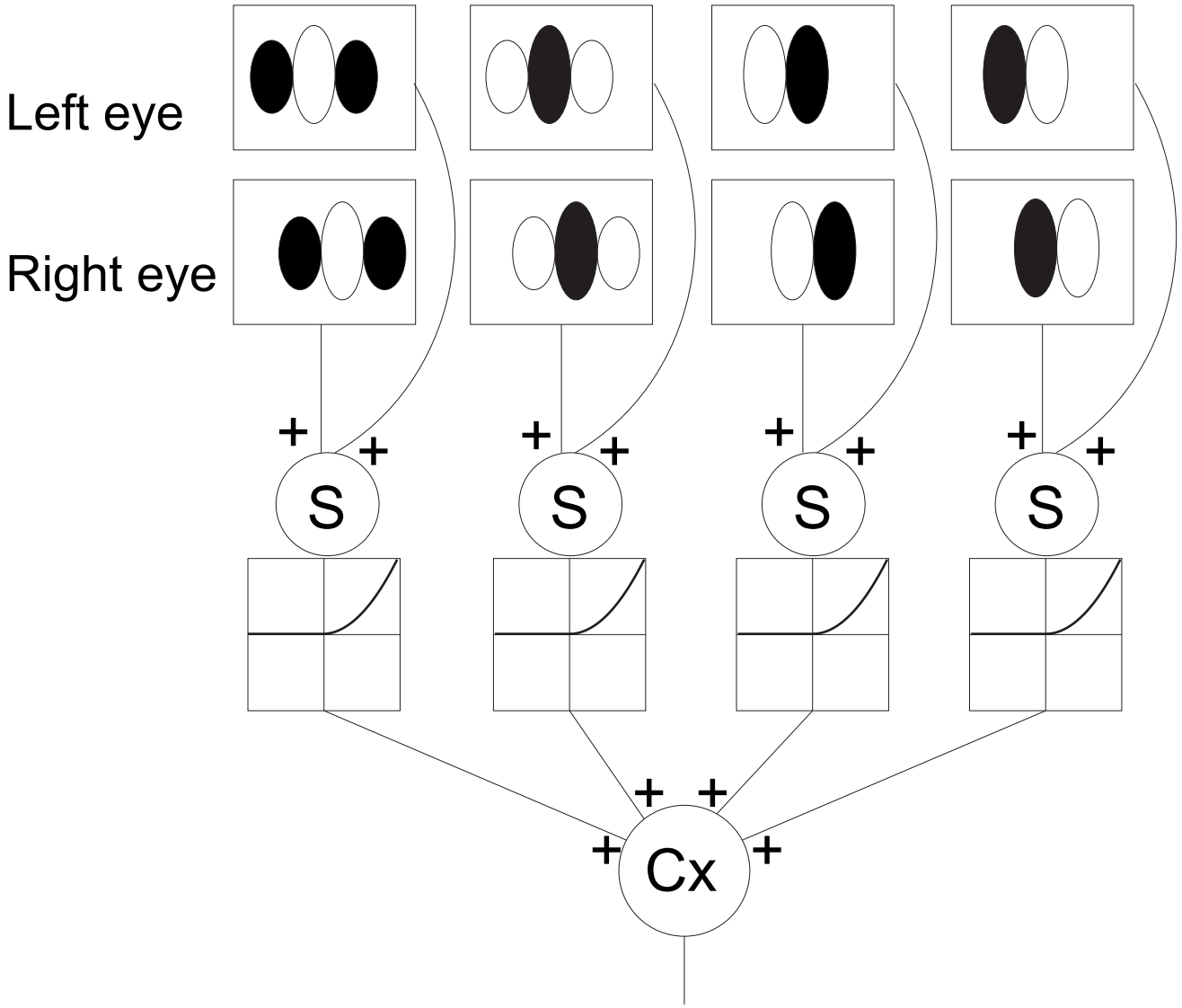

Figure 3. Structure of the disparity energy model proposed by Ohzawa et al. (1990) is illustrated. Four binocular simple-type subunits lying up in columns give inputs to a single complex-type unit. At the top, RFs for each binocular simple-type subunit are illustrated on the left and right retina (white squares). All subunits have a constant offset of RF position between the two retinae. The optimal orientation and SF is also common among subunits. Therefore, disparity encoded along the horizontal direction is identical across subunits. The phases of RFs of subunits are in quadrature pairs (a pair of the first column and the third column, and a pair of the second column and the fourth column). The phase of one subunit differs from the phase of the other by 90 ° in a pair. Inputs from the left and right retina are summed (indicated by plus symbol), half-wave rectified and squared (illustrated by a panel showing output as a function of summed binocular inputs). Outputs of these four subunits are simply summed into a single complex-type unit.

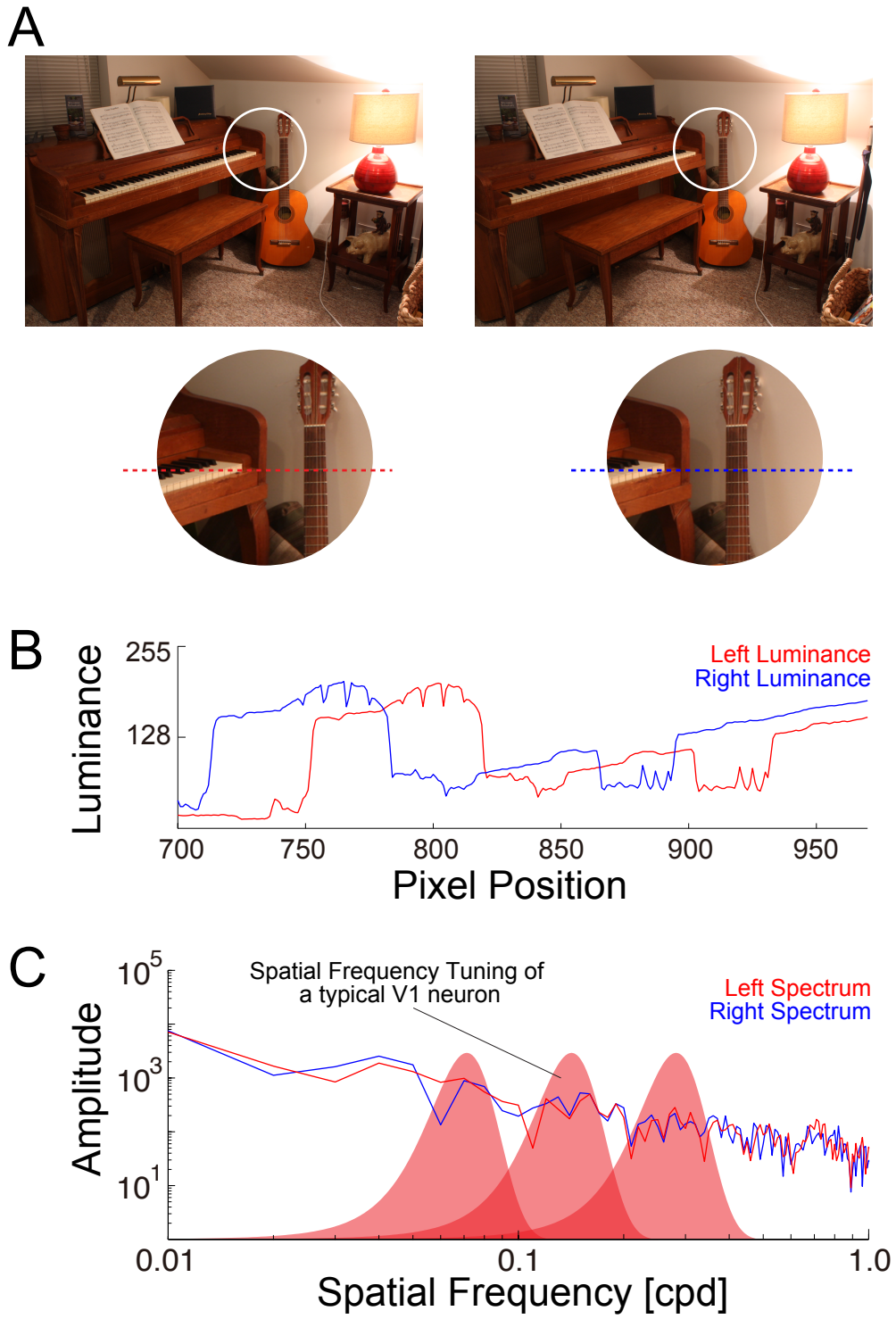

Figure 4. ***A***, A pair of stereo photographs is shown (Source: Middlebury Stereo Datasets). Small parts of the left and right images within white circles are enlarged for close examination. ***B***, Left and right luminance profiles for cross-sections of the images along the broken lines in A are plotted (red: left image, blue: right image). Cross-sections were taken after conversion to gray scale. ***C***, Amplitude spectra of the profiles in B are shown. SF tunings of three typical V1 neurons are illustrated schematically with red filled curves. Multiple V1 neurons, tuned to different frequencies, are needed to represent the broad stimulus spectra. For reliable stereoscopic depth detection, integration of signals from multiple frequency bands may be required.

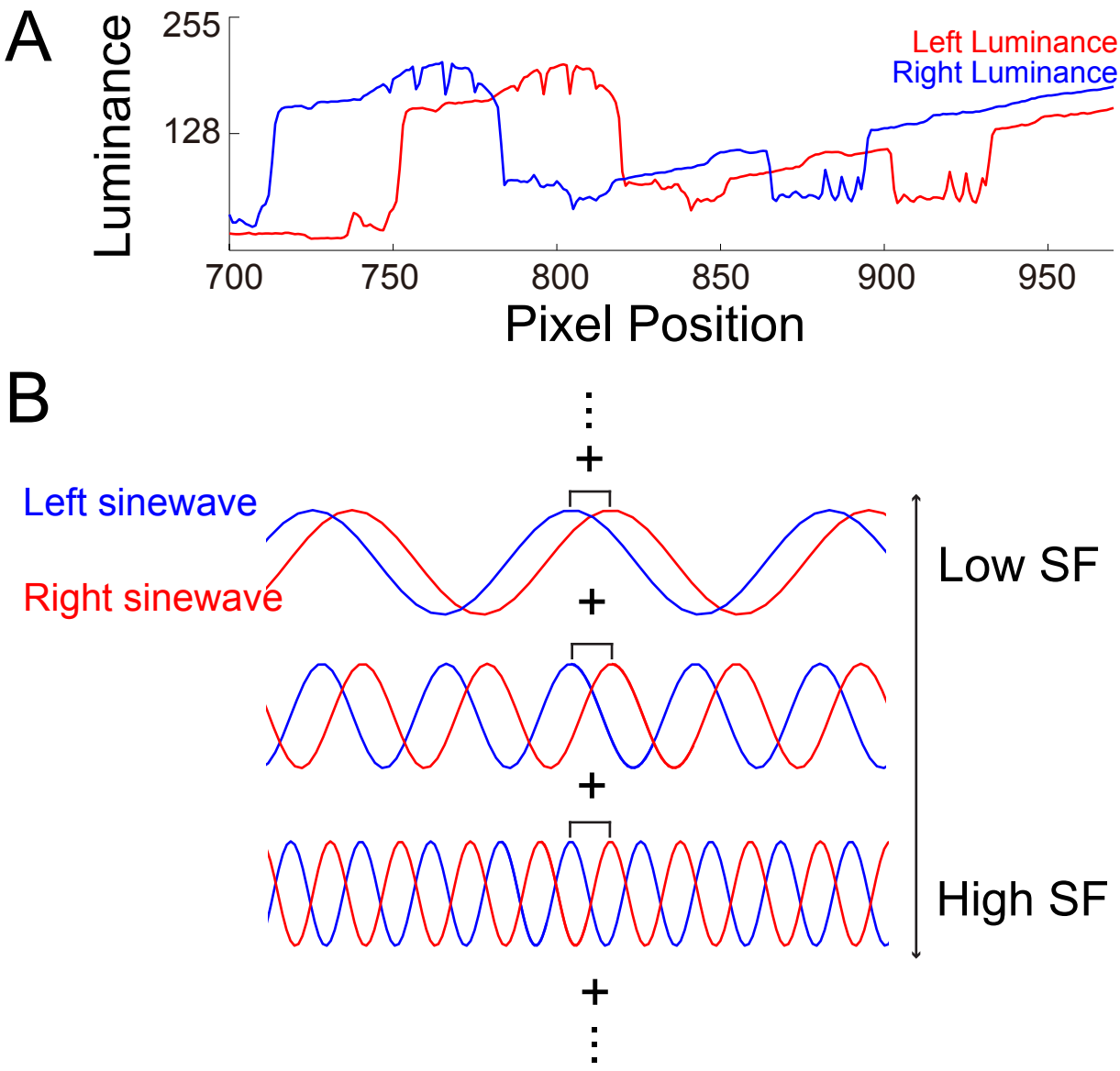

Figure 5. ***A***, Left and right luminance profiles that are the same as shown in Figure 4B are plotted (red: left image, blue: right image). ***B***, A complex luminance pattern is represented schematically as a combination of many sine-waves with different SFs. Binocular disparity, that is, the shift of luminance patterns between the two eyes as shown in A is equal to a constant amount of position shift between the left and right sine-waves (blue and red curves) across different SF components.

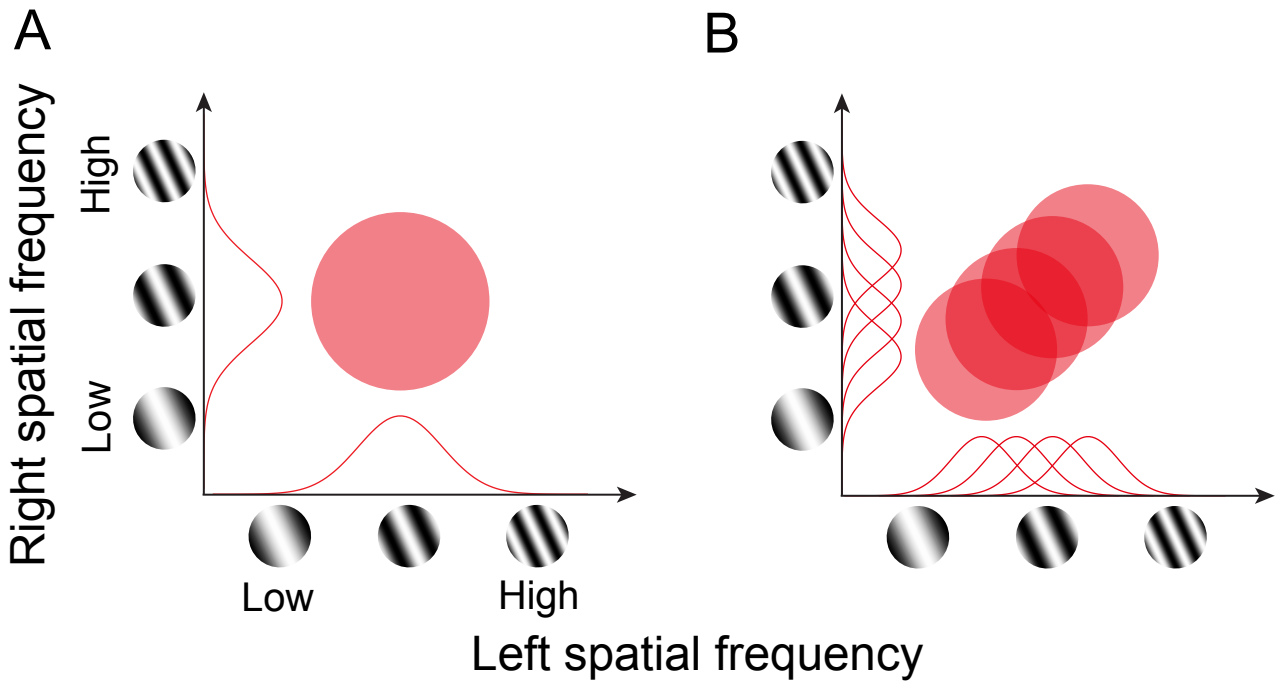

Figure 6. Predictions of binocular responses are illustrated schematically for neurons with and without pooling in the SF domain. ***A***, Neuron without pooling in the SF domain. A neuron having only a single SF channel would show a circular response profile in the joint left-right SF domain. Monocular SF tunings are indicated by red curves on horizontal and vertical axes for left and right eyes, respectively. ***B***, Predicted binocular response profile for a neuron with pooling in the SF domain. In this scheme, multiple subunits as shown in A but with different optimal SFs are pooled. The profile lies on the diagonal, because V1 neurons are generally tuned to similar SFs through either eye.

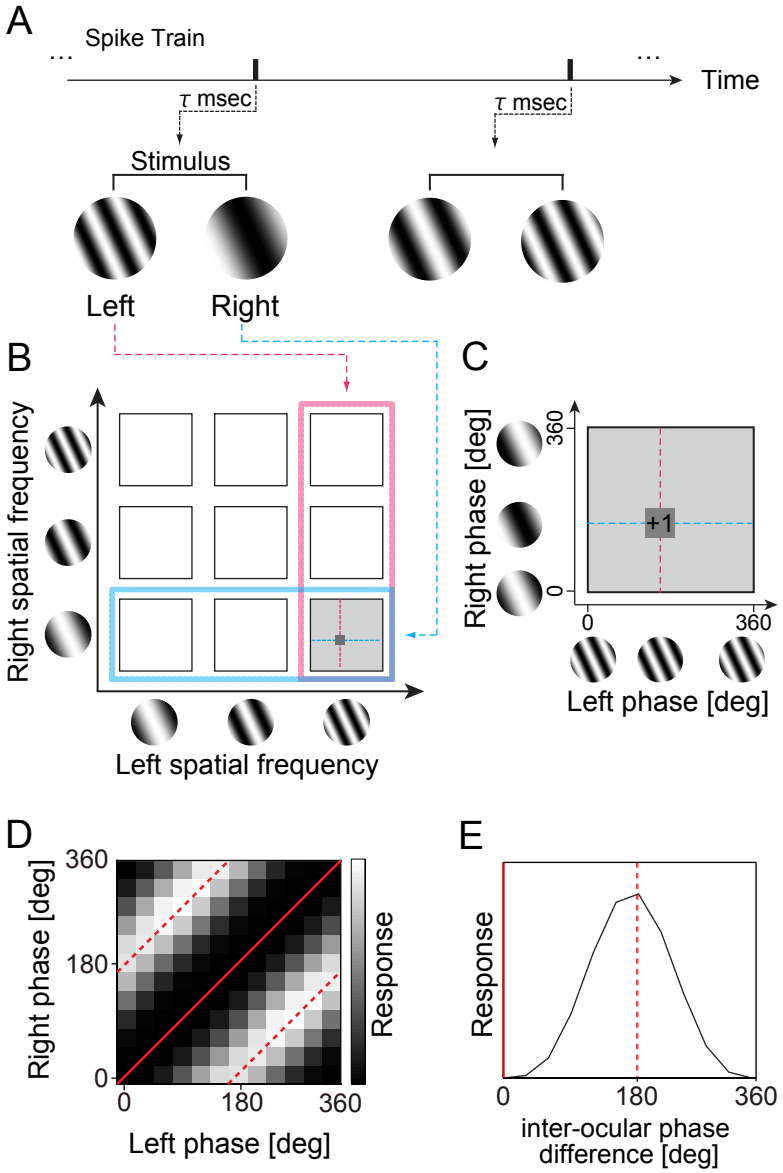

Figure 7. Reverse correlation analysis was performed to obtain binocular phase selectivity maps for various SF pairs. ***A***, Using spike data recorded while static sine-wave grating stimuli were presented dichoptically in rapid flashes, a pair of spike-triggered stimuli was selected for an optimal correlation time delay. The spike-triggered SF and phase parameters in the left and right eyes were extracted for the selected grating pair. ***B***, A spike was voted in a 4-dimensional domain (sf_L, sf_R, ph_L, ph_R), where "sf" and "ph" denote SF and phase, respectively, and the subscript indicates the eye. For clarity of understanding the 4-dimensional data, the procedure may be understood in the following steps. First, spike-triggered pair of SFs (sf_L, sf_R) was determined (left, pink; right, cyan). The selected pair is indicated by a gray square. ***C***, Then, within the selected joint phase subdomain (ph_L, ph_R), the spike was voted into a particular phase combination for the triggered stimulus pair. Repeating this procedure for all recorded spikes, complete binocular phase selectivity maps were obtained for various SF pairs between the two eyes. ***D***, The model response of a typical disparity energy unit is illustrated in the joint (ph_L, ph_R) subdomain, for the combination of optimal (sf_L, sf_R). The strength of response is shown with gray scale. Constant inter-ocular phase difference lines are indicated with red color (straight line for 0 degree and broken line for 180 degree). ***E***, One-dimensional tuning curve to inter-ocular phase difference was derived from D by integrating it along constant disparity lines.

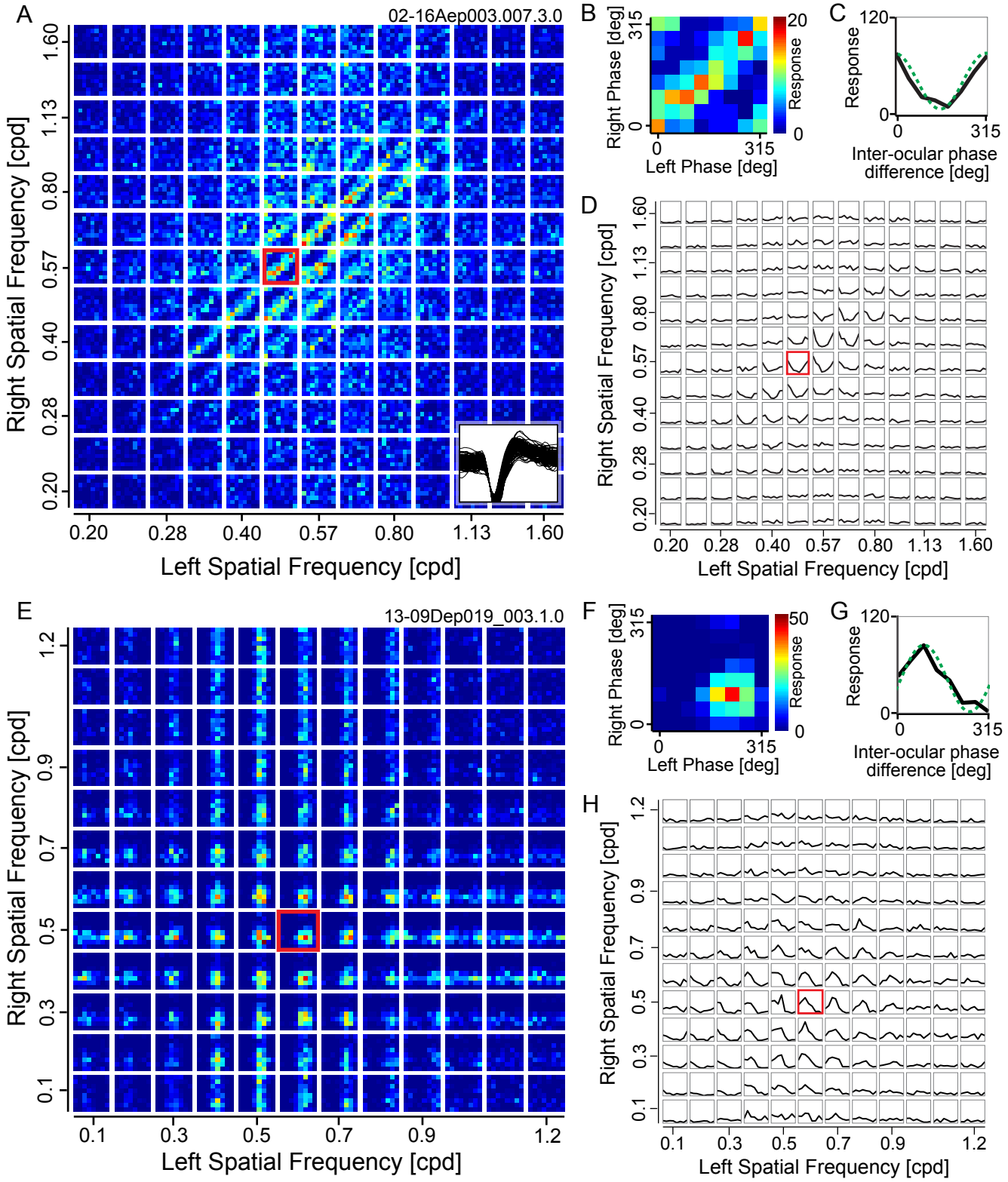

Figure 8. In A-D, responses are shown for a representative complex cell to stimuli containing various combinations of phases and SFs between the two eyes. ***A***, The entire data are response strengths in a 4-dimensional parameter space (sf_L, sf_R, ph_L, ph_R), and presented here as multiple cross sections (ph_L, ph_R) at each (sf_L, sf_R) for clarity. Each small domain is a binocular phase-selectivity map (ph_L, ph_R), illustrating responses to various combinations of left and right phases. These small maps are arranged as a 13x13 matrix of left and right SFs (sf_L, sf_R). Color scale shows the number of spikes collected for each pair of phases for the two eyes. One of the phase selectivity maps is marked with a red border and is magnified in B. Example spike waveforms of this cell is superimposed at the bottom-right corner (100 spikes). ***B***, Selectivity to binocular phase combination for the highlighted SF pair in A is shown. Strong responses are observed along a 45 degree diagonal where the inter-ocular phase difference was constant. ***C***, One-dimensional tuning curve to inter-ocular phase difference was calculated from a binocular phase selectivity map shown in B by integrating responses to the same inter-ocular phase difference between the two eyes. A sinusoid was fitted to the tuning curve to determine the modulation amplitude (green broken curve).

D, Tuning curves to inter-ocular phase difference are represented for all binocular combinations of SFs. The highlighted SF pair shown in C is indicated with a red border. The optimal SF and orientation for the dominant eye of this cell were 0.64 cpd and 19 degree, respectively. Values were similar for the non-dominant eye (0.61 cpd, 8 degree). In E-H, responses are shown for a representative simple cell to stimuli containing various combinations of phases and SFs between the two eyes. Panels ***E***, ***F***, ***G***, and ***H*** are shown in the same format as A, B, C, and D. Optimal SF and orientation for the dominant eye of this cell were 0.56 cpd and 9 degree, respectively. Values were similar for the non-dominant eye (0.53 cpd, 30 degree).

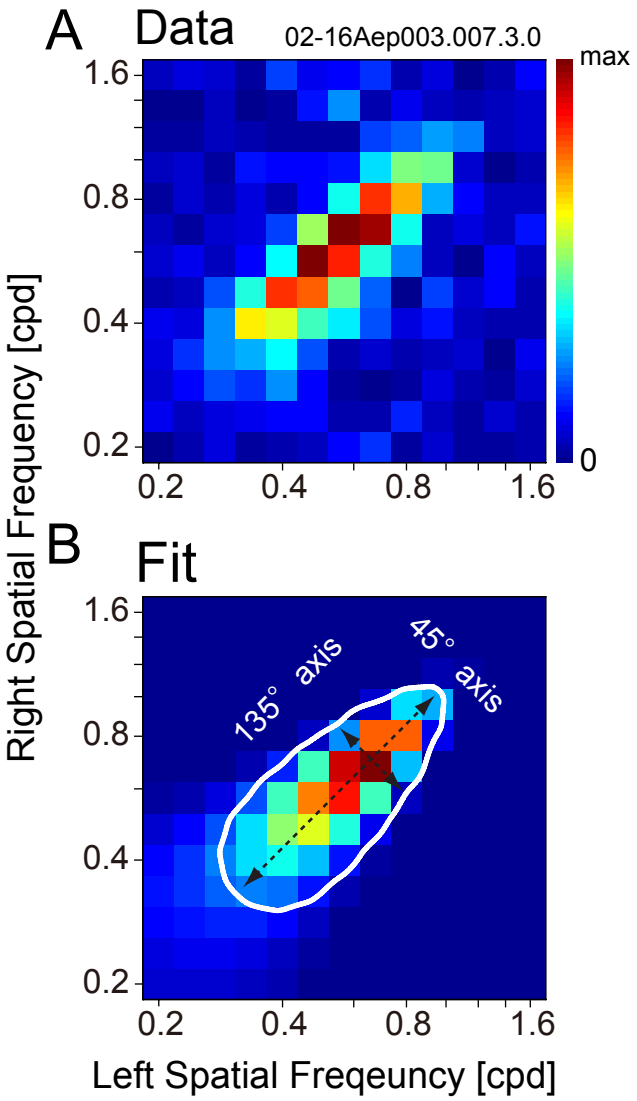

Figure 9. Binocular SF interaction maps were fitted for each neuron by a two-dimensional Gaussian function. ***A***, A binocular SF interaction map was obtained by extracting the degree of modulation (i.e., amplitude of F1 component) in tuning curves to inter-ocular phase difference for all binocular SF combinations (e.g., Fig. 4D). ***B***, The two-dimensional Gaussian function that yielded the best fit to the data presented in A is shown. The axes of a Gaussian were constrained to the direction of 45 degree and 135 degree diagonals. The sigmas of the fitted Gaussian along the 45 degree and 135 degree axes were extracted to calculate the index which indicates a degree of integration of SF channels. Specifically, elongation index in the SF domain (EIsf) was defined as a ratio of a sigma along the 45 degree axis to that along the 135 degree axis. EIsf for this cell was 3.13.

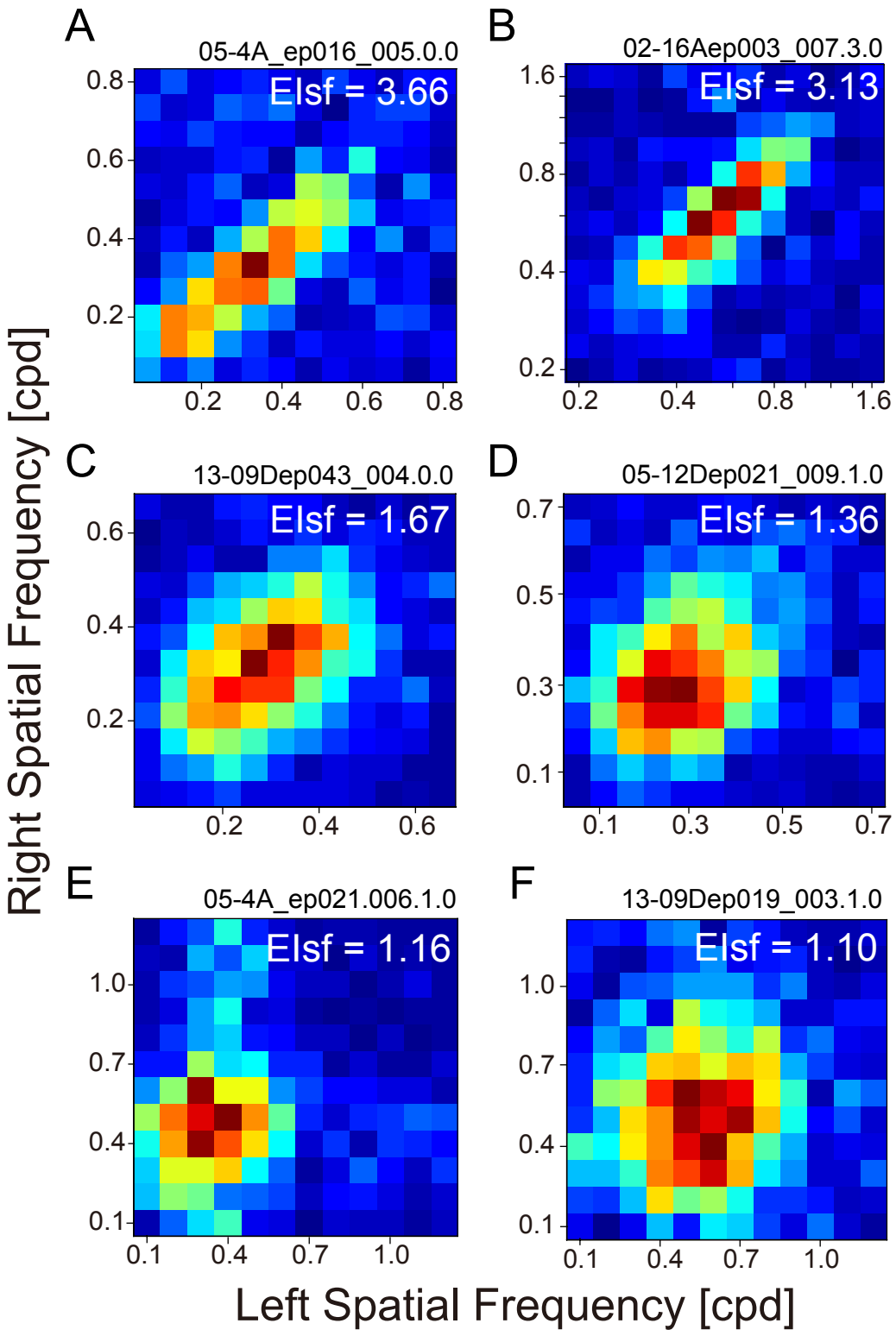

Figure 10. Binocular SF interaction maps showed various degrees of elongation across neurons. These maps are arranged in the descending order of EIsf. Value of EIsf is shown at the upper-right corner. ***A***, Optimal SF was 0.36 cpd, and optimal orientation (OR) was 163 degree for the dominant eye. ***B***, SF: 0.64 cpd, OR: 19 degree. ***C***, SF: 0.30 cpd, OR: 96 degree. ***D***, SF: 0.27 cpd, OR: 1 degree. ***E***, SF: 0.38 cpd, OR: 103 degree. ***F***, SF: 0.56 cpd, OR: 9 degree. Values were similar for the non-dominant eye (not shown). Cells for E and F were of simple type, while the remaining cells were complex.

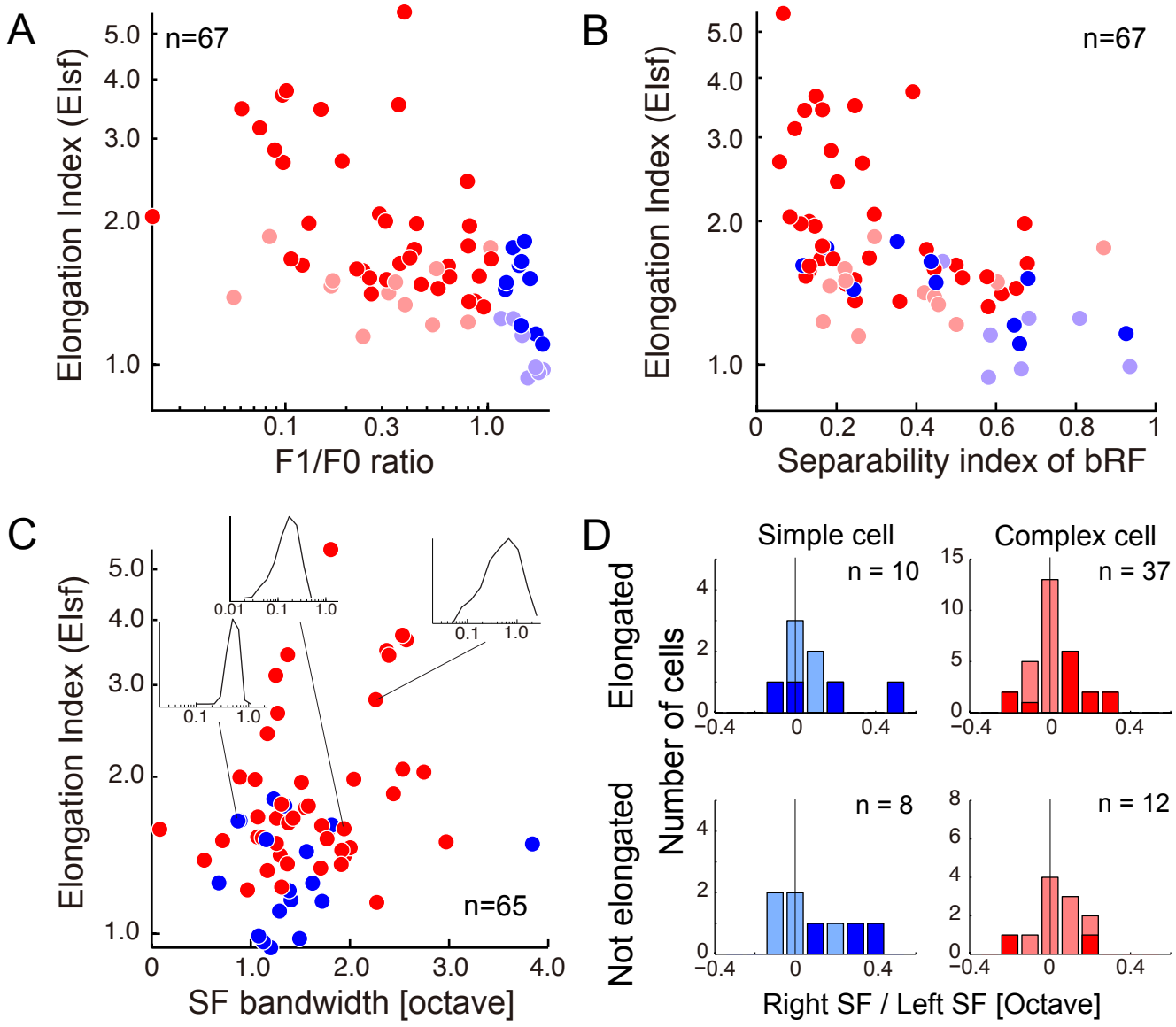

Figure 11. Possible relationships are examined between the EIsf (elongation of binocular SF map) and other response properties of cells. ***A***, Relationship between F1/F0 ratio and EIsf is presented as a scatter plot for all the recorded cells (n = 67). Negative correlation was observed between these two values ($r = -0.55$, $p < 0.001$). Blue: simple cells and red: complex cells. Deeper color indicates data with significant elongation (EIsf > 1.0, bootstrap test, $p < 0.01$). ***B***, Relationship between separability of binocular receptive field (bRF) and EIsf is presented as a scatter plot (n = 67). Detailed procedure of bRF calculation is explained in Figure 9. There was a significant negative correlation between these two values ($r = -0.61$, $p < 0.01$). ***C***, Relationship between SF bandwidth and EIsf is plotted similarly (n = 67). Insets show SF tuning curves measured with drifting sinusoidal grating stimuli for three representative cells. ***D***, Distributions of left and right optimal SF difference are summarized in the histograms. Horizontal axis shows optimal SF difference in an octave unit. The upper row shows neurons with significant elongation in the SF domain (EIsf > 1.0) while the lower shows those without elongation. Color usage is the same as other panels (Blue: simple cells and red: complex cells). Deeper color shows data with statistically significant SF difference obtained by a bootstrap test ($p < 0.01$).

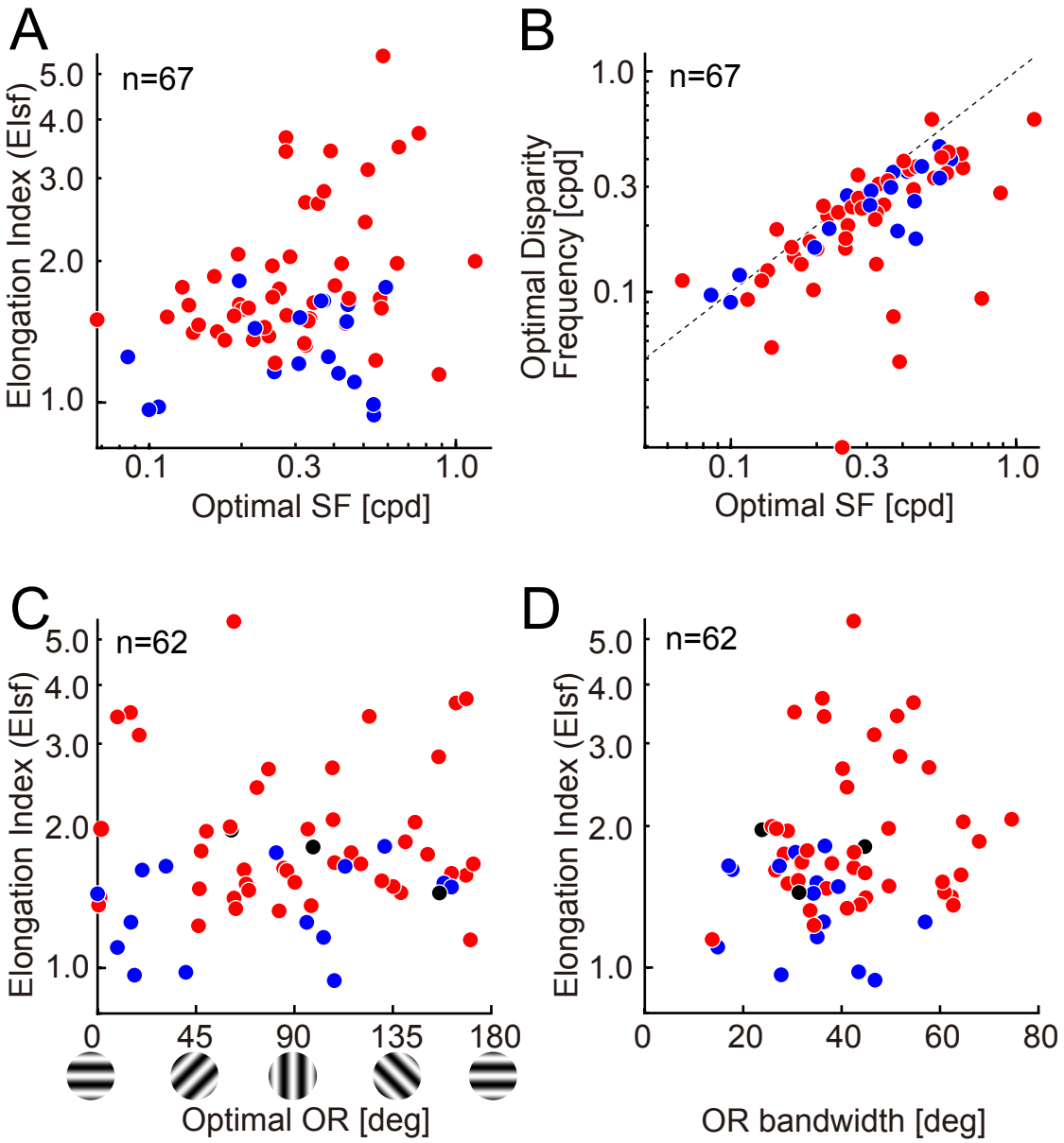

Figure 12. Relationships among basic parameters are presented as scatter plots for all the recorded cells. Blue: simple cells, Red: complex cells, and Black: unclassified (for C and D). ***A***, Relationship between optimal SF and EIsf is shown. Positive correlation was observed between these two values ($r = 0.24$, $p < 0.05$). ***B***, Relationship between optimal SF and optimal disparity frequency is shown. Broken black line indicates 1:1 identity line. Positive correlation was observed between the two values ($r = 0.73$, $p < 0.001$). Most cells showed lower optimal disparity frequency than monocular optimal SF. ***C***, Relationship between optimal orientation and EIsf is presented. Orientation here is defined that horizontal is 0 degree, and the value increases counterclockwise as illustrated at the bottom of the panel. ***D***, Relationship between orientation bandwidth and EIsf is presented.

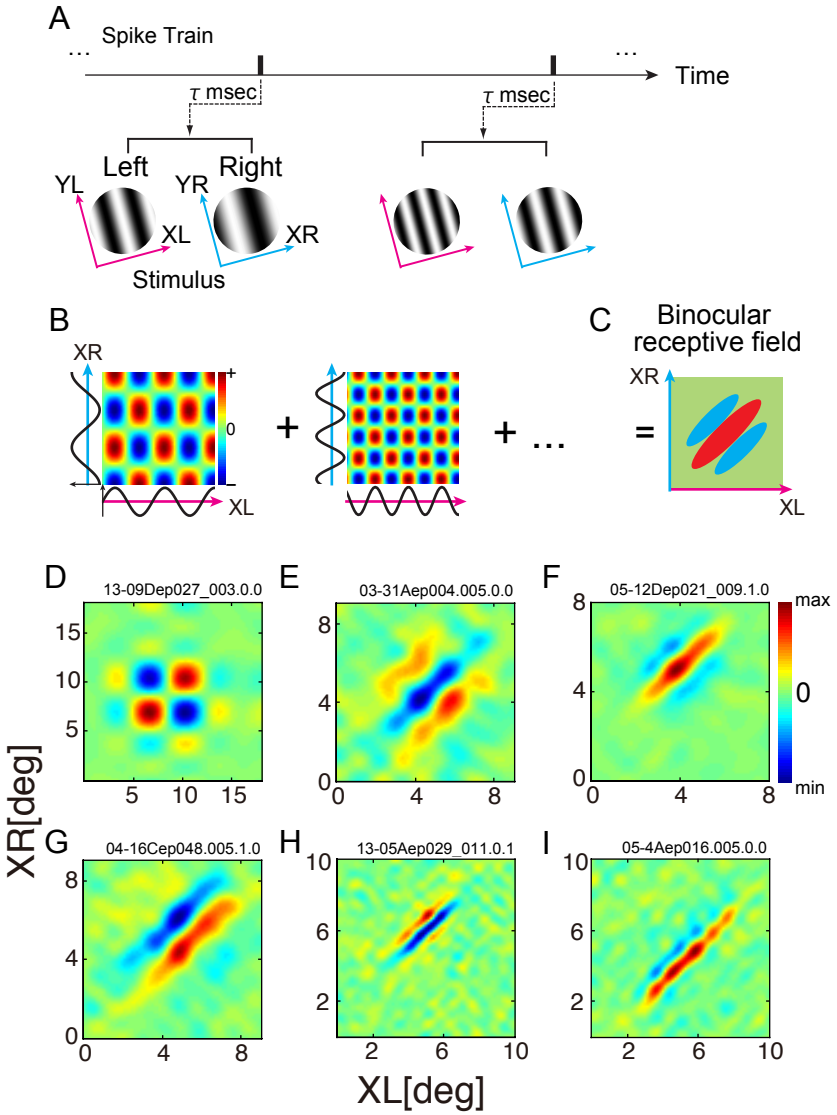

Figure 13. Although stimuli were defined in the SF and phase domain, it is possible to perform a reverse correlation analysis in the joint left-right space domain to obtain a binocular receptive field (bRF) for a given neuron. ***A***, Using a spike train recorded while binocular sine-wave grating stimuli were presented, spike-triggered stimulus pairs were selected for an optimal correlation time delay. ***B***, Binocular interaction terms were calculated from each pair of spike-triggered gratings as follows. The XL- and XR-axes are defined as the axis orthogonal to preferred orientation in the left and right eyes, respectively. These one-dimensional spike-triggered sinusoids in the XL- and XR-domains were multiplied to produce binocular interaction terms in the joint XL-XR domain. Positive values (red) mean that stimuli with the same contrast polarity were presented between the two eyes, whereas negative values (blue) mean those with the opposite polarity were presented between the two eyes. ***C***, A bRF was obtained by summing these interaction terms for all spike-triggered stimulus pairs. Binocular disparity is constant along the +45 degree diagonal in the map, while disparity changes along −45 degree diagonal. In D-I, reconstructed bRFs are shown for several representative cells. The horizontal axis of each map indicates position in XL-axis while the vertical axis indicates that in XR-axis. ***D***, BRF of a simple cell. Optimal SF, optimal orientation and EIsf of this cell was 0.09 cpd, 176 degree, and 0.97, respectively. All the remaining cells were of complex type, and the parameters are indicated in the above order. ***E***, Tuned inhibitory bRF. (SF: 0.25 cpd, OR: 73 degree, EIsf: 1.62) ***F***, Even-symmetric bRF. (SF: 0.27 cpd, OR: 1 degree, EIsf: 1.36) ***G***, Odd-symmetric bRF. (SF: 0.21 cpd, OR: 98 degree, EIsf: 1.81) ***H*** and ***I***, BRFs of cells that had relatively large EIsf. Parameters for the cell in H (SF: 0.40 cpd, OR: 20 degree, EIsf: 3.42), and I (SF: 0.36 cpd, OR: 163 degree, EIsf: 3.66). The bRFs allow evaluation of pooling along X axis (Sanada & Ohzawa, 2006) as EIx. Values of EIx were 1.04, 1.39, 1.80, 1.79, 3.39, 4.31, for D–I, respectively.

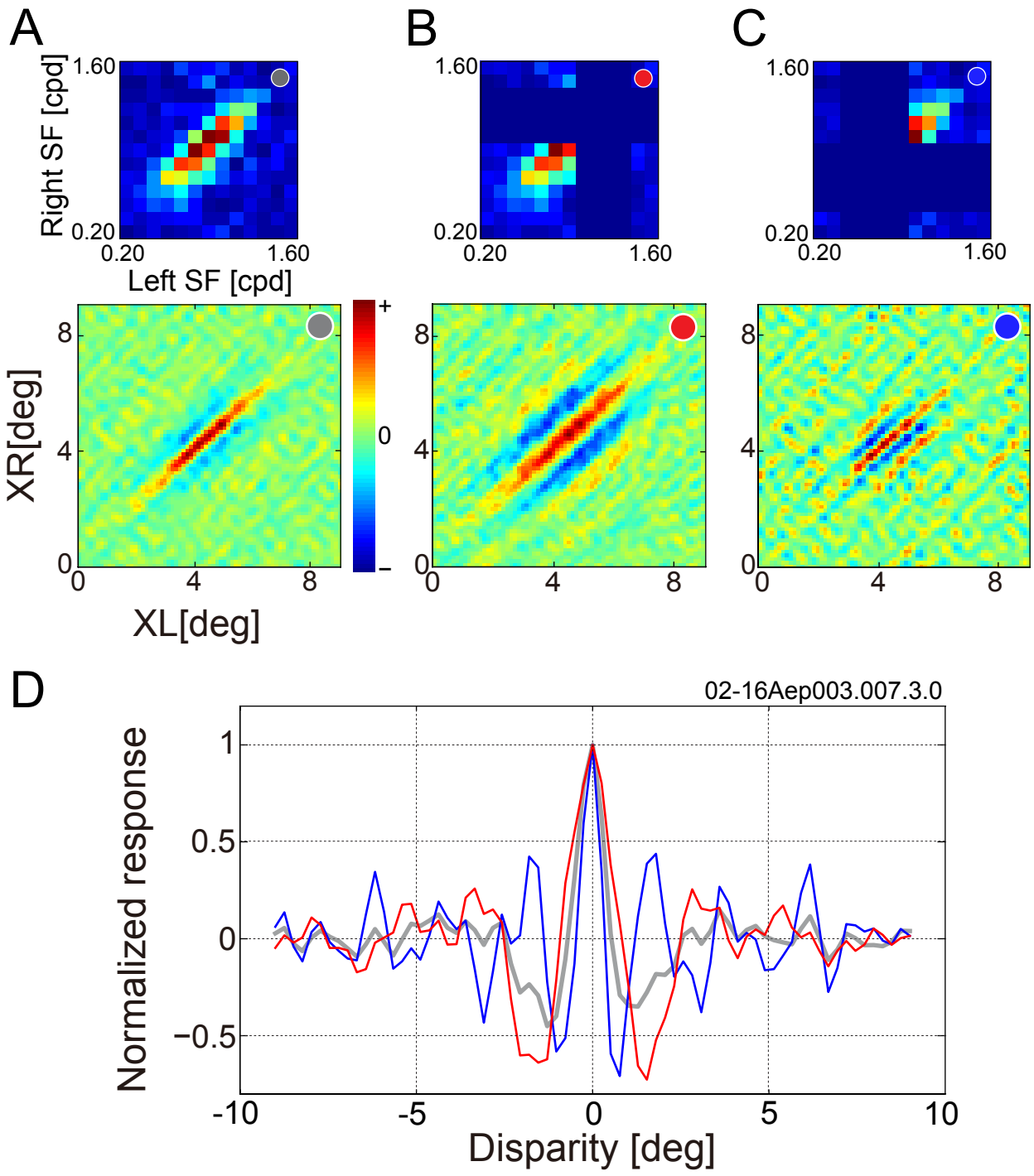

Figure 14. To allow visualization of properties of underlying pooled subunits, bRFs were reconstructed in different SF bands for a representative complex cell. In A-C, bRFs were reconstructed using limited SF bands. Binocular SF interaction maps are presented in the top row to show SF components used to reconstruct bRFs. Reconstructed bRFs are shown in the bottom row. ***A***, All SF components were used for reconstruction. ***B***, Only a lower half of the SF components was used for reconstruction. ***C***, Only an upper half of the SF components was used. Components that were not used for reconstruction are indicated by the darkest blue color. ***D***, Disparity tuning curves were obtained from the three bRFs. Color of the curves represents SF bands used for reconstruction (gray: all SF components, red: low SF components, and blue: high SF components).

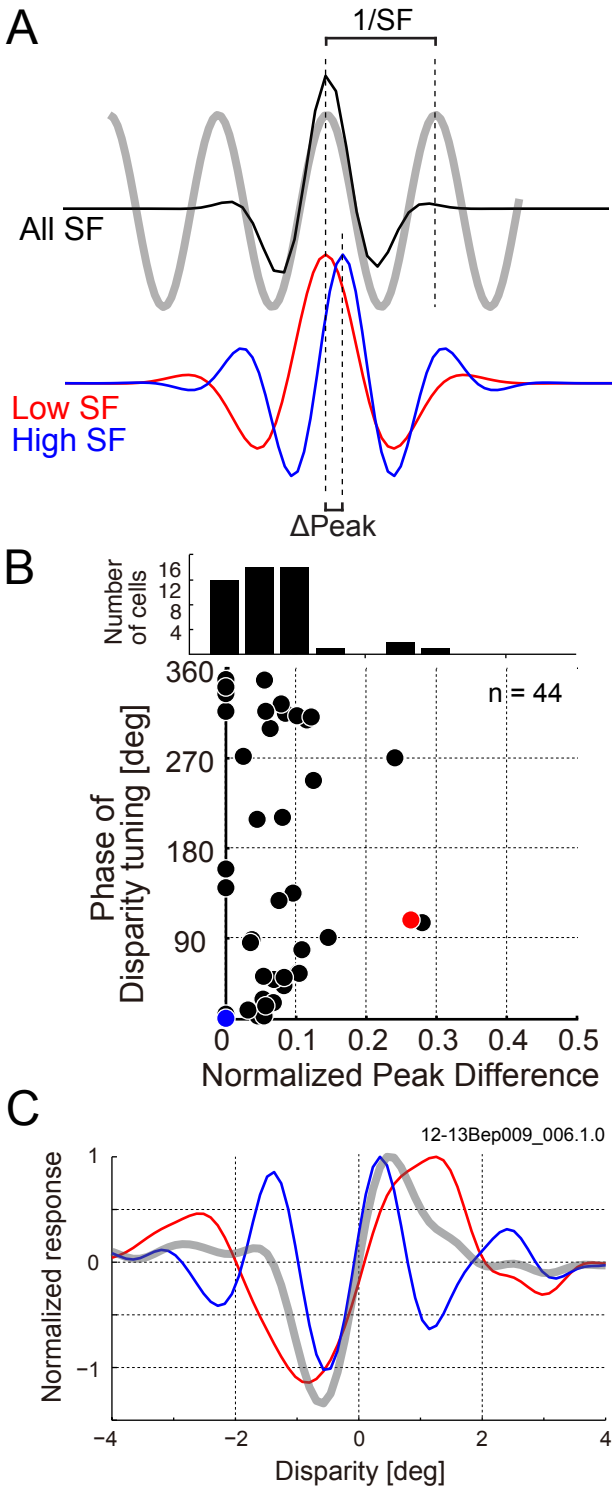

Figure 15. Are pooled SF sub-bands tuned to a common binocular disparity? ***A***, Alignment of preferred disparities is evaluated between low and high SF sub-bands by normalized peak difference. ΔPeak is the separation between the peaks originally in the unit of degrees. Normalized peak difference is the Δ Peak expressed as a fraction of the wavelength (1/SF), where SF is determined from the tuning curve using all frequency components. ***B***, Relationship between normalized peak difference and symmetry of disparity tuning curve (phase of fitted Gabor function) is plotted for 44 cells, which showed EIsf larger than 1.5. The blue dot indicate a neuron shown in Figure 10. The histogram on the top shows distribution of normalized peak difference. ***C***, An example cell that shows odd-symmetric tuning curve marked with red color in panel B is indicated. Disparity tuning curves obtained from three different bRFs in the same format as shown in Fig. 10 are superimposed, suggesting zero-crossing alignment across different SF bands (gray: all SF components, red: low SF components, and blue: high SF components).

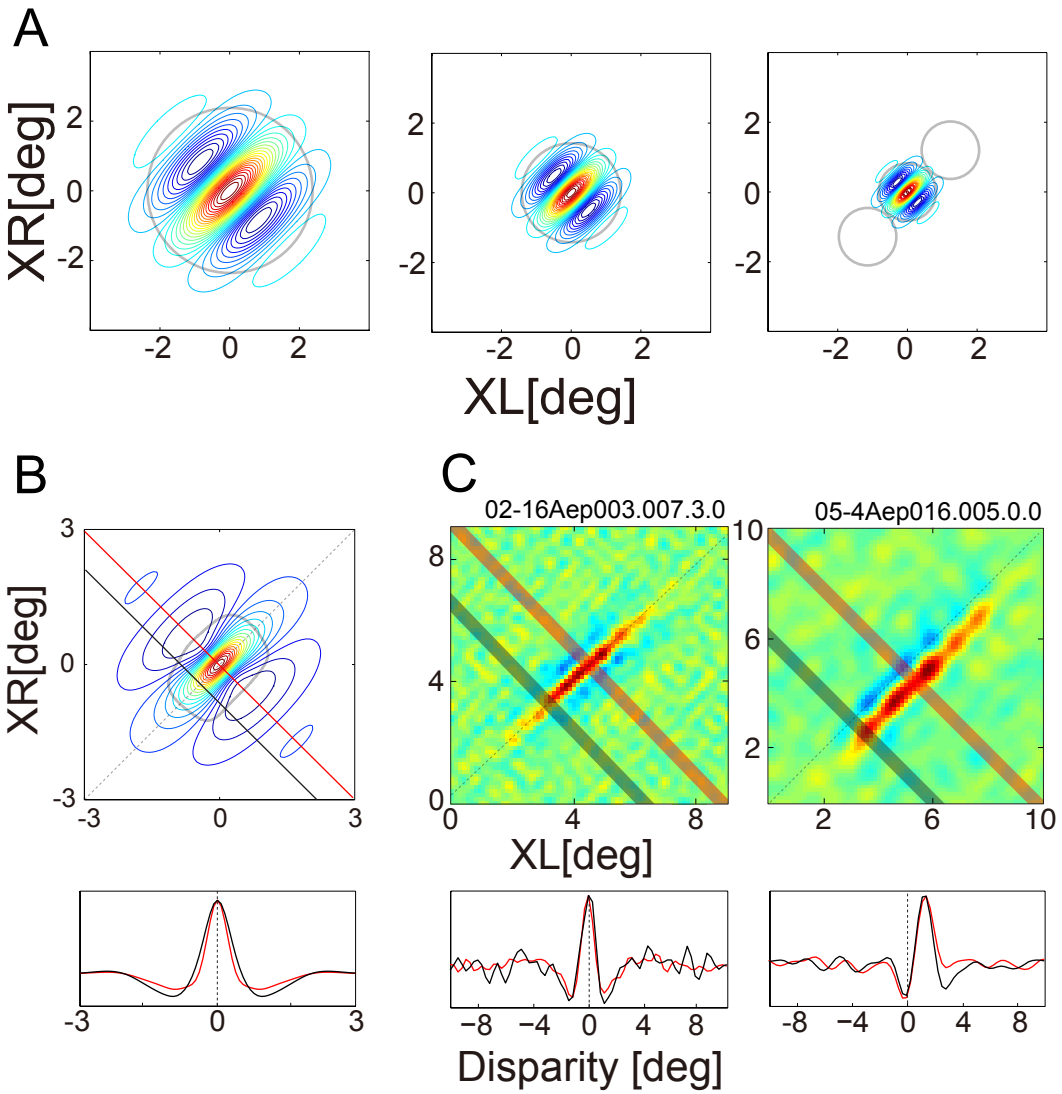

Figure 16. ***A***, Predicted response of three energy model units are shown in the format of bRFs. The three units are tuned to different SFs separated in steps of 1.7 (≈√3) times (therefore, the highest SF is 3 times larger than the lowest). Light gray circle indicates 10 % contour of the envelope. For the unit tuned to the highest SF, additional two locations are shown diagonally in the bRF space to indicate a possibility of greater degree of spatial pooling at higher SF. ***B***, A predicted response of a neuron that pools the three units shown in A is indicated. This is a model implementation of a neuron with pure SF pooling but without spatial pooling. The bRF is shown in the upper panel, with the envelope of bRF (25% of peak) depicted by a diagonally elongated light gray contour, calculated by Hilbert transform. Cross-sections of the bRF along the disparity axis (indicated as red and black lines) are compared in the bottom panel. Red curve shows a disparity tuning curve obtained as the cross-section at the center, while black curve shows the cross-section near the edge of the bRF (50% of the maximum amplitude of its envelope where there is little contribution from the highest SF unit). Each curve was normalized to be 1.0 at the maximum. ***C***, Data from two example neurons that had large EIx are shown. Upper panels show bRFs, while bottom panels show superimposed disparity tuning curves at the center and at the edge (50% of the maximum amplitude of its envelope) of each bRF. Each tuning curve was calculated as an average of cross-sections within band-shaped regions (shown as light red and black bands in the bRF).

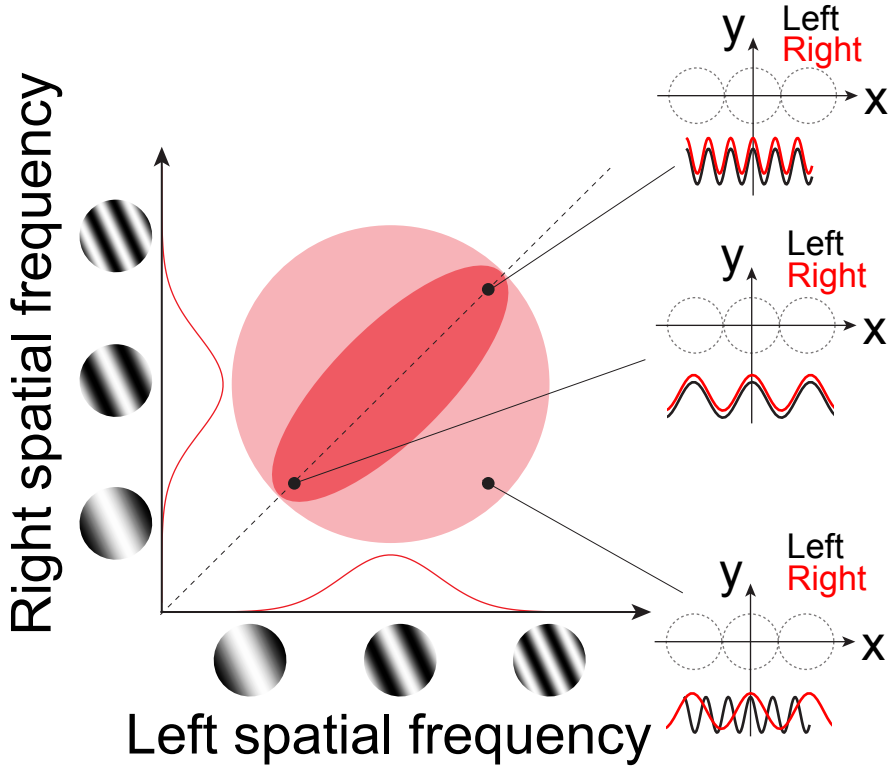

Figure 17. A schematic illustration explains how an apparent elongation of binocular SF interaction map may occur due to spatial pooling. In the left of the panel, a joint left-right SF tuning profile is illustrated as a light red circle for a neuron having a single SF channel. Each of three black dots on the response profile indicates different SF combination in the binocular SF domain. On the right, corresponding spatial arrangement of left and right sin-waves in the XY space domain is illustrated for each case. Broken gray circles show presumed spatial positions of different pooled subunits, and at the bottom of each row, cross-sections of left and right stimulus sin-waves are superimposed as black (left) and red (right) curves. When left and right SFs are the same between the two eyes like shown in the upper two cases, inter-ocular phase difference is always the same at any position along the X axis. However, in the case of different SF combinations between the two eyes, inter-ocular phase difference changes depending on the position along the X axis, generating different binocular phase arrangement between subunits at different positions. This causes mutual cancelling of inter-ocular phase tunings between subunits at different position by pooling them, resulting in apparent elongation of binocular SF interaction map along the 45° axis (red filled region).

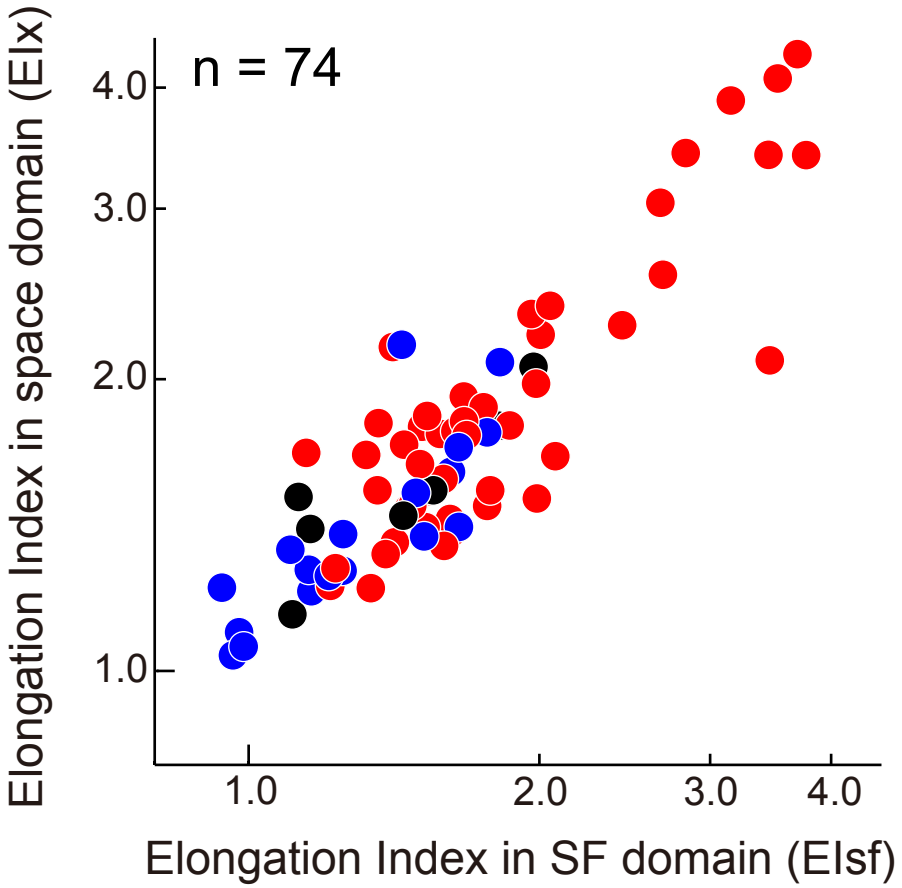

Figure 18. Relationship is shown between EIsf and elongation index in the space domain along the X-axis (diagonal) in the bRF (EIx). Symbol colors blue, red and black indicate simple, complex, and unclassified cells, respectively. A significant positive correlation was observed between the two values ($r = 0.8$, $p < 0.001$).

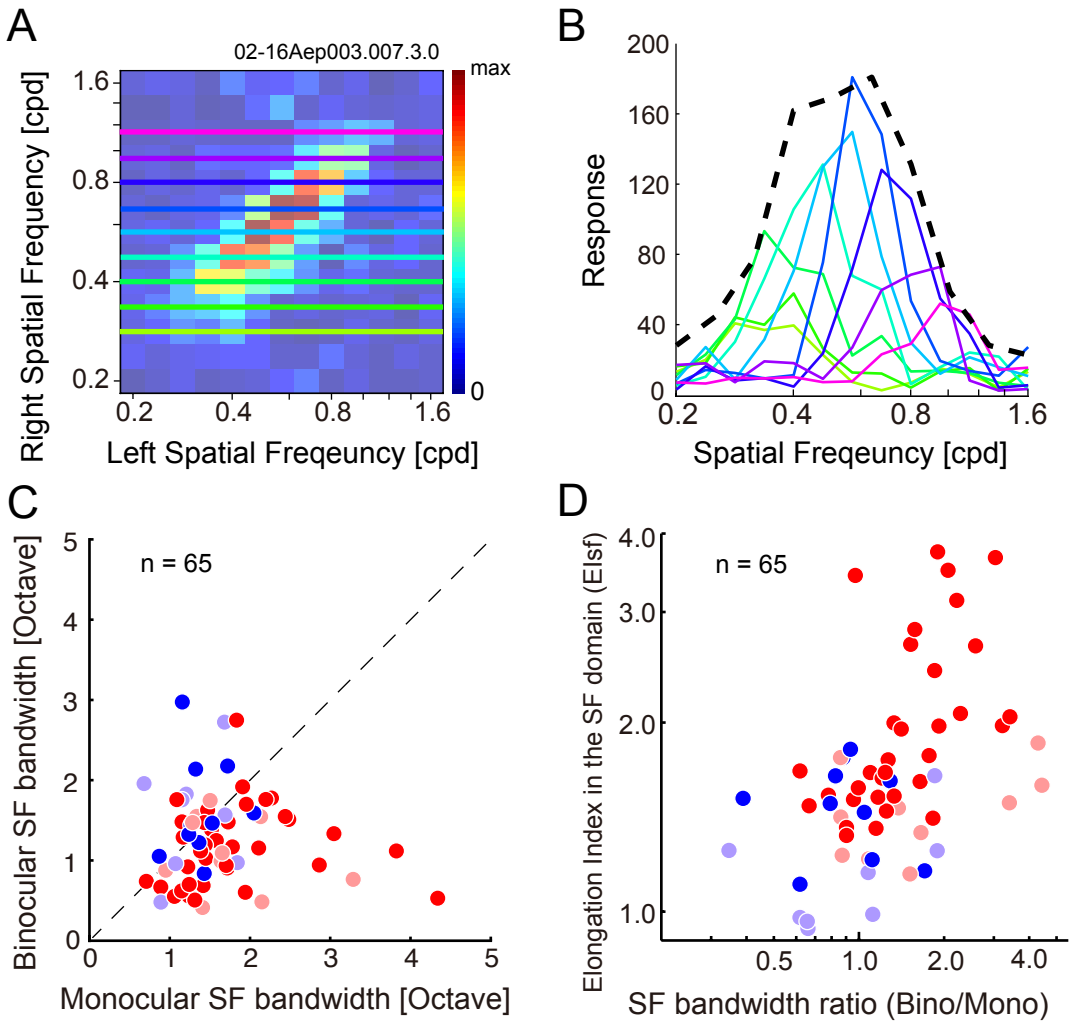

Figure 19. ***A***, To obtain binocular SF tunings, cross-sections were obtained in the binocular SF interaction map by fixing SF presented in the non-dominant eye at various values. Thus, binocular SF tuning curves were obtained for the dominant eye (in this case, left eye). The neuron shown in panel A and B is identical to one presented as a representative complex cell in Figure 8A. ***B***, Binocular SF tuning curves at various SFs in the right eye are superimposed for the neuron in A. Different color shows different SF in the right eye (as used in A). Monocular SF tuning obtained by drifting grating stimuli is indicated by a broken black curve. ***C***, Comparison of SF tuning bandwidth between monocular and binocular conditions is shown for a population of neurons. Among many binocular SF tuning curves (as presented in B), the tuning curve showing the maximum response (i.e., at the peak of a binocular interaction map) was used for binocular SF tuning. Broken black line indicates identity line between the bandwidth of two conditions. Color usage of the dots is the same as previous Figures (Blue: simple cells, Red: complex cells). ***D***, EIsf is compared with the discrepancy of SF bandwidth between monocular and binocular conditions (SF bandwidth ratio). SF bandwidth ratio for each neuron was obtained from values shown in C.

業績リスト

原著論文

Baba M, Sasaki KS, Ohzawa I. (2015) Integration of multiple spatial frequency channels in disparity-sensitive neurons in the primary visual cortex. J Neurosci in press.

国内学会

M. Baba, K.S. Sasaki, S. Man, I. Ohzawa. Comparison of binocular and monocular neural response time course in primary visual cortex of cats. 第 34 回日本神経科学大会. 2011/9/16

M. Baba, K.S. Sasaki, I. Ohzawa. Analysis of V1 disparity-sensitive neurons in the spatial frequency-phase difference domain. 第 36 回日本神経科学大会. 2013/6/21

M. Baba, K.S. Sasaki, I. Ohzawa. Integration of multiple spatial frequency channels in V1 disparity detectors. 第 37 回日本神経科学大会. 2014/9/12

国際学会

M. Baba, K.S. Sasaki, I. Ohzawa. V1 complex cells pool multiple disparity detectors tuned to different spatial frequencies. Neuroscience 2013. 2013/11/10

M. Baba, K.S. Sasaki, I. Ohzawa. Integration of multiple spatial frequency channels in V1 disparity detectors. 新学術領域研究「質感脳情報学」国際シンポジウム "Future of Shitsukan Research". 2014/7/16,17

M. Baba, K.S. Sasaki, I. Ohzawa. Integration of multiple spatial frequency channels in V1 disparity detectors. The 10th Asia-Pacific Conference on Vision. 2014/7/21

受賞等

The 10th Asia-Pacific Conference on Vision にて Best Student Presentation Award 受賞