



# Joint-encoding of motion and depth by visual cortical neurons: neural basis of the Pulfrich effect

Akiyuki Anzai<sup>1,2</sup>, Izumi Ohzawa<sup>1,3</sup> and Ralph D. Freeman<sup>1</sup>

<sup>1</sup> Group in Vision Science, School of Optometry, University of California, Berkeley, California 94720-2020, USA

<sup>2</sup> Present address: Department of Anatomy and Neurobiology, Washington University School of Medicine, St. Louis, Missouri 63110, USA

<sup>3</sup> Present address: Department of Biophysical Engineering, Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka, 560-8531, Japan

Correspondence should be addressed to R.D.F. ([freeman@neurovision.berkeley.edu](mailto:freeman@neurovision.berkeley.edu))

Motion and stereoscopic depth are fundamental parameters of the structural analysis of visual scenes. Because they are defined by a difference in object position, either over time or across the eyes, a common neural machinery may be used for encoding these attributes. To examine this idea, we analyzed responses of binocular complex cells in the cat striate cortex to stimuli of various intra- and interocular spatial and temporal shifts. We found that most neurons exhibit space–time-oriented response profiles in both monocular and binocular domains. This indicates that these neurons encode motion and depth jointly, and it explains phenomena such as the Pulfrich effect. We also found that the relationship between neuronal tuning of motion and depth conforms to that predicted by the use of motion parallax as a depth cue. These results demonstrate a joint-encoding of motion and depth at an early cortical stage.

Motion is characterized by a series of positions an object traverses over time. An object presented sequentially at two positions, with a time delay between them, is sufficient for us to perceive motion<sup>1</sup>. Likewise, depth can be signaled by two positions, one in each eye; a difference in position between images in left and right eyes (binocular disparity) is sufficient to produce our perception of depth<sup>2</sup>. Therefore, the extraction of motion and depth from retinal images involves similar processes that could use a common encoding mechanism. Because the striate cortex is the first site along the central visual pathways at which neural selectivities for direction of motion and binocular disparity arise, neurons in this area may be involved in joint-encoding of motion and depth. However, the encoding of motion and that of binocular disparity have been studied separately, and the issue of joint-encoding has not been addressed.

Here we establish that single neurons in the striate cortex encode motion and binocular disparity jointly. This joint-encoding is of direct relevance to a compelling and intriguing visual illusion, the Pulfrich effect, which is described below. For this reason, we first examine the neural basis of the Pulfrich effect to illustrate the joint-encoding. We then determine a relationship between the tuning properties of individual neurons for motion and binocular disparity. This relationship is critical for determining whether joint-encoding has an advantage over independent-encoding.

A pendulum swinging back and forth along a plane perpendicular to an observer's line of sight seems to follow an elliptical path in depth when seen binocularly with a light-attenuating filter in front of one eye<sup>3</sup>. This phenomenon, the Pulfrich effect, has been investigated extensively in psychophysics<sup>4–6</sup> but minimally in physiology<sup>7,8</sup>. It is known that the filter reduces the transmission speed of neural signals in the filtered eye<sup>7–11</sup>, thereby creating an interocular difference in signal latency. This in

turn changes the perceived depth of the pendulum<sup>11</sup>. However, the neural mechanism that translates the interocular latency difference into the depth change is a matter of debate<sup>12</sup>.

The classical explanation of the Pulfrich effect is as follows<sup>3</sup>. Because the neural signal from the filtered eye reaches the visual cortex with a slight delay compared to the signal from the unfiltered eye, the pendulum position indicated by the signal in the filtered eye falls slightly behind that in the other eye at a given moment. This creates a spatial offset between the pendulum positions in the two eyes (binocular disparity), and shifts the apparent depth of the pendulum (Fig. 1a). One possible mechanism that underlies this process is a set of neurons tuned to various interocular spatial offsets specified by signals arriving simultaneously from the two eyes (interocular spatial offset hypothesis, Fig. 2a). This explanation is essentially based on simple stimulus geometry (Fig. 1a), rather than a neural mechanism that converts a delay in one eye to a change in the perceived depth. Although it accounts for the standard Pulfrich effect, it falls short of explaining Pulfrich-like phenomena such as those seen in dynamic noise patterns containing no coherent motion<sup>13,14</sup> and stroboscopic stimuli with no interocular spatial offset<sup>5,6,15,16</sup>. This suggests that a specific physiological process underlies the Pulfrich effect and its related phenomena.

An alternative explanation is that the visual system uses the latency difference between the neural signals originating from corresponding points on the two retinas to compute the depth of the pendulum relative to the fixation depth<sup>6,17</sup>. An object moving either in front of or behind the fixation plane stimulates a point on the retina in one eye first and then its corresponding point in the other eye (Fig. 1b). Therefore, the latency difference between the neural signals from corresponding points may indicate a moving object

**Fig. 1.** The Pulfrich effect. **(a)** A classical explanation of the Pulfrich effect. A light-attenuating filter (gray rectangle) placed in front of one eye causes a time delay for the neural signal in that eye. Therefore, the perceived position of the swinging pendulum in the filtered eye (gray circle) falls slightly behind that in the other eye (open circle) for a given moment, creating a spatial offset between the perceived images in the two eyes. Left, when a pendulum swings to the right while the right eye is covered with a filter, the spatial offset corresponds to a binocular disparity that shifts the perceived depth of the pendulum toward the observer (hatched circle). Right, when the pendulum swings to the left, the spatial offset corresponds to a binocular disparity that shifts the perceived depth of the pendulum away from the observer (hatched circle). The amount of the binocular disparity depends on the pendulum position along the fronto-parallel plane because the speed of the pendulum motion varies with the position. Therefore, the pendulum appears to move along an elliptical trajectory (solid arrow) in depth as it swings back and forth on the fronto-parallel plane (dashed arrow). **(b)** A time delay between stimulation of the corresponding points in the two eyes by a moving object. An object moving to the right in front of the fixation plane F (or to the left behind the fixation plane) stimulates a retinal point in the left eye ( $P_L$ ) first (at  $t_1$ ) and then, with some time delay (at  $t_2$ ), the corresponding point in the right eye ( $P_R$ ). Therefore, a latency difference between the neural signals from the corresponding points in the two eyes may be used as a cue for a moving object not on the fixation plane. **(c)** Space–time trajectories of an object moving along a constant depth plane at a constant speed. The two solid lines indicate retinal positions of left and right eye images of the object over time, respectively. The horizontal distance between the two lines corresponds to the binocular disparity of the object, and the tilt of the lines from vertical indicates speed of the object motion. Interocular spatial offset and time difference can be specified between any combination of two points on the two lines. For instance, the object position in the right eye at time  $t_2$  can have interocular spatial offsets of  $D_1$ ,  $D_2$  and  $D_3$ , and interocular time differences of  $\Delta T_1$ ,  $\Delta T_2$  and  $\Delta T_3$ , with each of the object positions in the left eye at time  $t_1$ ,  $t_2$  and  $t_3$ , respectively. There is a linear relationship between the interocular spatial offsets and time differences (inset).

not on the fixation plane. However, interocular latency difference alone is not sufficient to determine the relative depth of the object; for a given latency difference, the magnitude of the relative depth depends on the speed of object motion, and the sign of the relative depth is determined by the direction of object motion. To explain the Pulfrich effect with this scheme, the visual system requires mechanisms tuned to various interocular latency differences at corresponding points (that is, at zero interocular spatial offset), as well as a mechanism that converts the latency differences into depth changes according to the speed and direction of the pendulum motion (interocular time difference hypothesis, Fig. 2b).

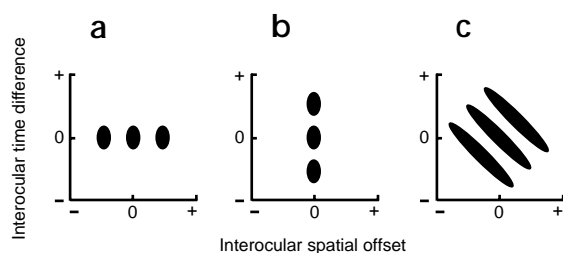
There is also a third possibility. Interocular spatial offset and latency difference may both contribute to the determination of the relative depth of the pendulum<sup>6,16,18</sup>. Consider an object moving at a given depth. Left and right eye images of the object traverse the retinas along the trajectories illustrated by the two solid lines in Fig. 1c. Horizontal and vertical distances between the two lines correspond to the interocular spatial offset and time difference, respectively. Instead of detecting one of the distances for a given space or time (as in the previous two hypotheses), the visual system may integrate measurements of interocular spatial offset and time difference for various combinations of two points along the trajectories. Such a strategy is expected to produce a robust signal of depth, because it uses information gathered over a range of space and time. Because there is a linear relationship between the interocular spatial offset and time difference (Fig. 1c, inset), this hypothesis predicts that neurons that carry

out this process exhibit a similar linear relationship in their tuning functions (space–time integration hypothesis, Fig. 2c).

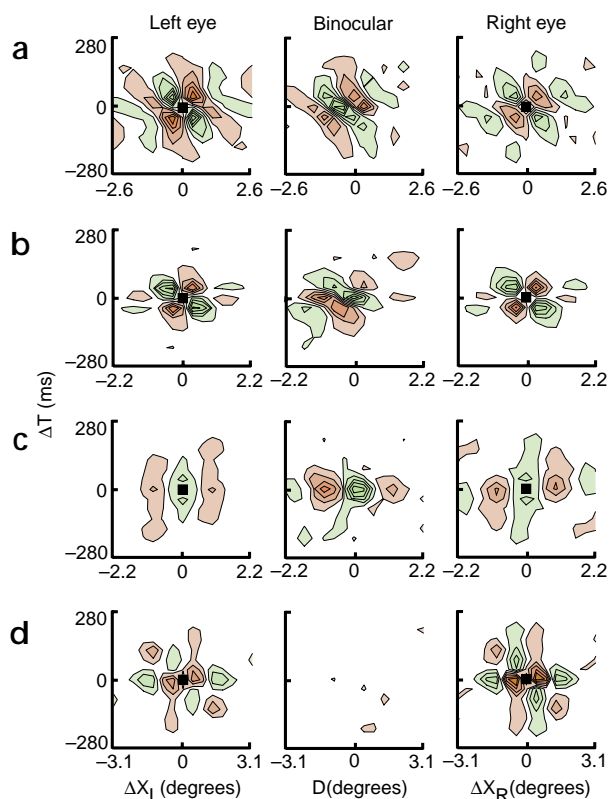
To test these hypotheses, we used an efficient receptive field mapping technique<sup>19</sup> and examined responses of neurons in the cat's striate cortex for various interocular spatial offsets and time differences of stimuli. We found that most neurons exhibit tuning that is consistent with the space–time integration hypothesis (Fig. 2c). This supports the idea that motion and depth are encoded jointly. We also found that these neurons are tuned to specific combinations of stimulus speed and binocular disparity, suggesting that the joint-encoding carries information unavailable to independent-encoding of motion and depth. These results indicate that motion and depth are encoded jointly at the initial stage of cortical processing.

## RESULTS

We measured responses of single neurons in the cat striate cortex to dichoptic presentations of one-dimensional noise patterns that consisted of 16 bars in each eye<sup>20,21</sup>. From the responses, we extracted components due to interactions between any combination of two bars. The interactions were then examined according to the spatial offset ( $\Delta X$ ) and time difference ( $\Delta T$ ) between the two bars, for cases in which the two bars belonged to the same eye (monocular interaction) or where the two bars were imaged in opposite eyes (binocular interaction). These interactions characterize how neurons respond to motion and binocular disparity ( $D$ ); interactions between the two bars that appear sequentially at different loca-



**Fig. 2.** Hypothetical tuning profiles of neurons for interocular spatial offset and time difference. Each oval indicates a region of responses for a single neuron. **(a)** Interocular spatial offset hypothesis. Neurons are tuned to an interocular time difference of zero, but to various interocular spatial offsets. **(b)** Interocular time difference hypothesis. Neurons are tuned to an interocular spatial offset of zero, but to various interocular time differences. **(c)** Space–time integration hypothesis. Neurons are tuned to various combinations of interocular spatial offset and time difference in such a way that there is a linear relationship between them.



**Fig. 3.** Examples of two-bar interaction profiles. Profiles of four binocular complex cells (**a–d**) are shown for the left eye (left column), right eye (right column) and binocular (middle column) domains. The  $x$ - and  $y$ -axes of each profile indicate the relative spatial offset ( $\Delta X_L$ ,  $\Delta X_R$  and  $D$ ) and time difference ( $\Delta T$ ) between two bars in the stimulus, respectively. When two bars belong to different eyes, the  $x$ - and  $y$ -axes represent an interocular spatial offset (that is, binocular disparity,  $D$ ) and an interocular time difference between the two bars, respectively. Because an interaction of a bar with itself cannot be obtained with the stimulus used in this study (binary m-sequence noise), one data point at  $\Delta X_L$  (or  $\Delta X_R$ ) = 0 and  $\Delta T$  = 0 in each of the monocular profiles is missing and has been filled with a black square. If this data point were available, the two green contour regions in the second and fourth quadrants of the maps would have been connected through the point to form an elongated region (as has been shown previously<sup>22</sup>), just like those in the binocular profiles. Green and red contour regions indicate positive and negative values, respectively. For each cell, contour lines are drawn such that they divide the response amplitude between 0 and the greater of the positive and negative peaks of the three profiles into 6 equally spaced levels.

tions describe responses to motion, whereas those between the two bars that appear across the two eyes represent responses to binocular disparity. Our focus here is on binocular complex cells, because they are generally selective for binocular disparity and are insensitive to monocular spatial phase, so that their responses depend on relative, not absolute, positions of the two bars<sup>21</sup>.

Most of the binocular complex cells we examined (52/59) exhibited both monocular and binocular interactions (Fig. 3a–c). However, we also found, for a small number of cells (7/59), strong monocular but minimal binocular interactions (Fig. 3d). Because such cells cannot encode binocular disparity, they were excluded from the rest of our analysis. We never observed a binocular cell that exhibited binocular but not monocular interactions.

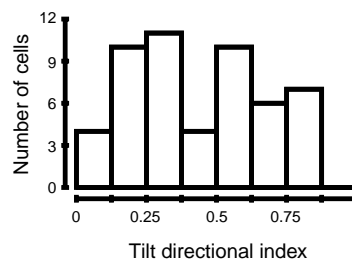
As previously reported<sup>22</sup>, many monocular profiles were obliquely oriented in the  $\Delta X - \Delta T$  domain (Fig. 3a and b, left and right columns), indicating that these cells were selective for the direction of stimulus motion. Binocular profiles of these cells were similarly oriented in the  $D - \Delta T$  domain (Fig. 3a and b, middle column). The tilt angle of the profiles varied from cell to cell, but for a given cell, monocular and binocular profiles were generally similar. (See below for a parametric analysis.) To quantify the degree of tilt in the binocular profile, we computed a tilt directional index (TDI) for each cell that described a response bias toward one direction of tilt with respect to the opposite direction of tilt (see Methods for definition of TDI). A TDI of zero indicated no bias, that is, there

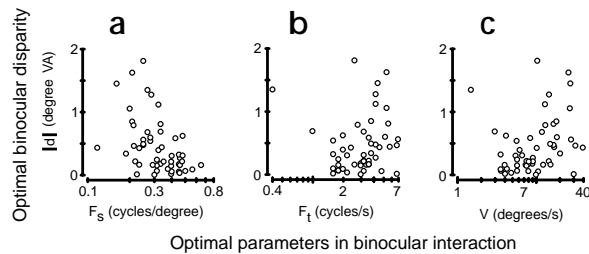
was no tilt, whereas a TDI of one meant the response was completely accounted for by one direction of tilt. The index is equivalent to the direction selectivity index used to quantify the response bias toward the cell's preferred direction of stimulus motion relative to the non-preferred direction<sup>23</sup>, and is analogous to the index that describes the degree of inseparability in space-time receptive fields of simple cells<sup>24</sup>. Indices for cells shown in Fig. 3a, b and c are 0.869, 0.839 and 0.076, respectively. The distribution of TDI for our population of cells ( $n = 52$ ) was widely spread with a mean of 0.44 (Fig. 4). Therefore, there was a clear bias for the binocular profile to be tilted from vertical.

A horizontal slice of the binocular profile represents a cell's tuning for binocular disparity, and, therefore, the obliquely oriented profiles indicated that disparity tuning shifted with interocular time difference of the two bars. Previously, it was shown that placing a light-attenuating filter in front of one eye causes a shift in tuning of neurons in the cat's striate cortex for interocular spatial phase disparity<sup>8</sup>. This result and ours presented here provide an explanation as to why the incorporation of an interocular latency difference into the neural signals causes a shift of the apparent depth of the pendulum in the Pulfrich effect. Whereas it could not be determined in the previous study<sup>8</sup> if the neurons were tuned to interocular spatial offset, interocular time difference or both, our results demonstrate clearly that it is both. The approximately linear relationship between interocular spatial offset and time difference, seen in the profiles, is consistent with the space-time integration hypothesis (Fig. 2c).

We also found that some cells exhibited profiles that were not oriented (Fig. 3c, TDI = 0.076). Such profiles seemed to be consistent with the prediction of the interocular spatial offset hypothesis (Fig. 2a). However, they were more likely to be part of a continuum of degree of tilt. This interpretation was consistent with the finding that monocular profiles also come in various amounts

**Fig. 4.** A population distribution of tilt directional index for binocular profiles. The distribution has a mean of 0.44 ( $n = 52$ ). An index of 0.25 means that the response to the preferred direction of tilt is greater than that to the non-preferred direction of tilt by 67%. Because more than 70% of the cells have an index of that magnitude or greater, there is a clear bias for the binocular profile to be tilted from vertical.





**Fig. 5.** Relationships between motion and depth tuning. Magnitude of the optimal binocular disparity ( $d$ ) in degree visual angle (degree VA) is plotted as a function of optimal spatial and temporal frequencies and speed for a population of cells ( $n = 52$ ). (a) Dependency on optimal spatial frequency. The range of binocular disparity decreases with spatial frequency. (b) Dependency on optimal temporal frequency. The range of binocular disparity increases with temporal frequency. (c) Dependency on optimal speed. The range of binocular disparity increases with speed. These results indicate that fine disparities are encoded by neurons tuned to various spatial and temporal frequencies and speeds, whereas coarse binocular disparities are encoded by neurons tuned to low spatial frequencies, high temporal frequencies and high speeds.

of tilt, depending on the cell's selectivity for the direction of stimulus motion<sup>22</sup>.

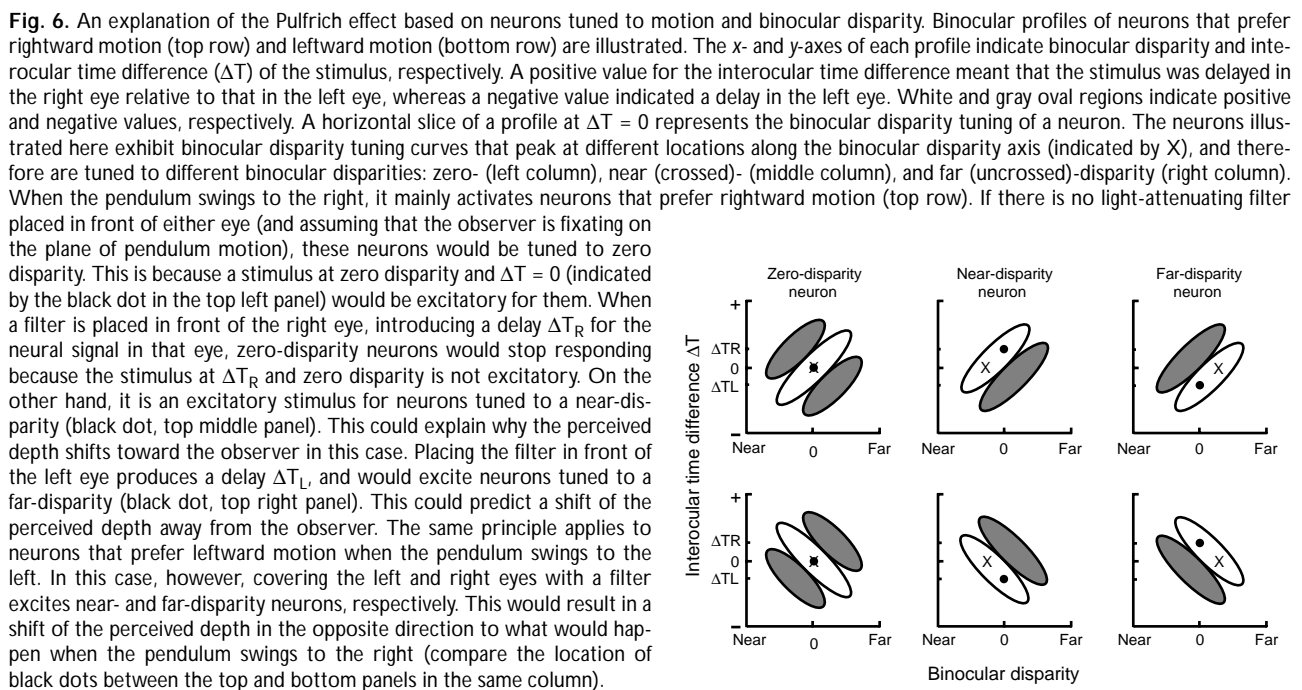
In a previous study<sup>7</sup> that examined the tuning of area 18 neurons in cats to interocular spatial offset and time difference, it was found that some cells responded best to non-zero interocular time differences. This result is consistent with the interocular time difference hypothesis (Fig. 2b). However, because details of the tuning were not examined, it was not possible to distinguish between different hypotheses based on the data. In fact, we found no cells that showed profiles consistent with the interocular time difference hypothesis.

The oriented profiles for monocular interactions are ideal for detecting moving stimuli because they match expected responses to such stimuli<sup>22</sup>. The oriented profiles for the binocular interactions shown here may be interpreted analogously. Human observers perceive motion in certain stimuli containing an interocular spatial offset and a time difference<sup>25–27</sup>, a phenomenon called dichoptic motion. The binocular profiles in the middle column of Fig. 3a and b are consistent with expected responses to dichoptic motion stimuli. Therefore, these neurons may constitute a neural basis for dichoptic motion as well. With regard to underlying mechanisms for dichoptic motion, there has been a debate as to whether early motion sensors are monocular or binocular<sup>28,29</sup>. Our results suggest that they are binocular and are tuned to binocular disparity.

These findings demonstrate that most neurons encode motion and binocular disparity jointly. Is there a relationship between the motion and binocular disparity encoded by individual neurons? If

there is, it may indicate that the visual system selects certain combinations of motion and binocular disparity to extract information that would not have been available had they been processed separately. On the other hand, if there is no correlation between them, the joint-encoding at this stage seems only to preserve information about concurrency of motion and binocular disparity. To address this question, we next examined the relationship between motion and binocular disparity tuning of neurons.

For the parameter that represents motion tuning, we estimated the optimal speed ( $V$ ) as a ratio of the optimal temporal frequency ( $F_t$ ) to the optimal spatial frequency ( $F_s$ ) for each profile. These three parameters were generally matched between monocular and binocular interactions for each cell. However, there was a small tendency for binocular optimal spatial frequency to be lower and binocular optimal temporal frequency to be slightly higher than the monocular counterpart. As a consequence, optimal speed tended to be slightly higher for binocular compared to monocular interaction ( $t$ -test;  $p < 0.01$ ). For the parameter that represented binocular disparity tuning, we estimated optimal binocular disparity ( $d$ ) for each binocular profile. The magnitude of optimal binocular disparity depended on the optimal spatial and temporal frequencies and the optimal speed of the binocular profiles (Fig. 5). We found that the range of binocular disparities encoded by a population of cells decreased with optimal



**Fig. 6.** An explanation of the Pulfrich effect based on neurons tuned to motion and binocular disparity. Binocular profiles of neurons that prefer rightward motion (top row) and leftward motion (bottom row) are illustrated. The  $x$ - and  $y$ -axes of each profile indicate binocular disparity and interocular time difference ( $\Delta T$ ) of the stimulus, respectively. A positive value for the interocular time difference meant that the stimulus was delayed in the right eye relative to that in the left eye, whereas a negative value indicated a delay in the left eye. White and gray oval regions indicate positive and negative values, respectively. A horizontal slice of a profile at  $\Delta T = 0$  represents the binocular disparity tuning of a neuron. The neurons illustrated here exhibit binocular disparity tuning curves that peak at different locations along the binocular disparity axis (indicated by  $X$ ), and therefore are tuned to different binocular disparities: zero- (left column), near (crossed)- (middle column), and far (uncrossed)-disparity (right column). When the pendulum swings to the right, it mainly activates neurons that prefer rightward motion (top row). If there is no light-attenuating filter placed in front of either eye (and assuming that the observer is fixating on the plane of pendulum motion), these neurons would be tuned to zero disparity. This is because a stimulus at zero disparity and  $\Delta T = 0$  (indicated by the black dot in the top left panel) would be excitatory for them. When a filter is placed in front of the right eye, introducing a delay  $\Delta T_R$  for the neural signal in that eye, zero-disparity neurons would stop responding because the stimulus at  $\Delta T_R$  and zero disparity is not excitatory. On the other hand, it is an excitatory stimulus for neurons tuned to a near-disparity (black dot, top middle panel). This could explain why the perceived depth shifts toward the observer in this case. Placing the filter in front of the left eye produces a delay  $\Delta T_L$ , and would excite neurons tuned to a far-disparity (black dot, top right panel). This could predict a shift of the perceived depth away from the observer. The same principle applies to neurons that prefer leftward motion when the pendulum swings to the left. In this case, however, covering the left and right eyes with a filter excites near- and far-disparity neurons, respectively. This would result in a shift of the perceived depth in the opposite direction to what would happen when the pendulum swings to the right (compare the location of black dots between the top and bottom panels in the same column).



spatial frequency (Fig. 5a), but it increased with optimal temporal frequency (Fig. 5b). As a result, the disparity range increased with optimal speed (Fig. 5c). In other words, fine disparities were encoded by neurons tuned to various spatial and temporal frequencies and speeds, whereas coarse disparities were encoded by neurons tuned to low spatial frequencies, high temporal frequencies, and therefore, high speeds. These results are in accord with human psychophysical data that indicate that stereoacuity can be maintained at higher speeds for sinusoidal gratings of low spatial frequencies compared with gratings of high spatial frequencies<sup>30</sup>, and that the upper disparity limit for coarse stereopsis increases with stimulus speed<sup>31</sup>.

#### DISCUSSION

Motion and binocular disparity constitute two distinct cues for the visual system to obtain spatial and temporal structures of a scene. Yet they are similar in that both can be described by interactions between two points in space and time. Therefore, it is likely that the visual system encodes them similarly. Our current results demonstrate that motion and binocular disparity are encoded jointly by single neurons in the striate cortex.

We showed that most complex cells in the cat striate cortex exhibited responses due to two-bar interactions for both motion and binocular disparity. Space-time oriented profiles of monocular interaction (Fig. 3a and b) represent motion and are known as motion energy responses, because they correspond to responses predicted by a mechanism that detects energy in a motion stimulus (a motion energy model<sup>32,33</sup>). A similar mechanism has been proposed to explain responses due to interocular two-bar interactions when there is no time difference between the two bars<sup>34,35</sup>. This mechanism is ideally suited for encoding binocular disparity and is called a binocular disparity energy model. The binocular profiles described in this paper, which include interocular time differences between the two bars, can easily be explained by incorporating a motion energy model into a binocular disparity energy model (a hybrid energy model<sup>18</sup>). Although a binocular disparity energy model itself does not constrain binocular profiles to have a certain shape in the space-time domain (except at  $\Delta T = 0$ ), a motion energy model provides a space-time-oriented structure to the profiles.

The structures seen in two-bar interaction, including the space-time oriented ones, have been interpreted as receptive field structures of simple-cell-like subunits that underlie individual complex cells<sup>21,22,36-38</sup>. Subunits are thought to combine inputs linearly over space and time<sup>21,33-36,39</sup>. This summation is linear regardless of whether it occurs within one eye or across the two eyes. Therefore, binocular interactions exhibit a structure similar to that of monocular interactions as long as the latter are similar for the two eyes. In other words, if monocular profiles are space-time oriented, so is the binocular profile (except for some complex cells such as that shown in Fig. 3d). In this sense, the space-time oriented structure of the binocular profile results from linear binocular subunits that are selective for the direction of stimulus motion. It is this space-time oriented binocular interaction that seems to be the underlying neural basis for the Pulfrich effect and dichoptic motion.

In light of our findings, the Pulfrich effect can be explained as follows. When the pendulum swings to the right, it mainly activates neurons that prefer rightward motion at around the speed of the pendulum (Fig. 6, top row). These neurons exhibit two-bar interaction profiles that are oriented to the right in the space-time domain. If there is no light-attenuating filter placed in front of either eye, then these neurons are the ones that are tuned to zero disparity (Fig. 6, top left). Now suppose that the right eye is covered with a filter, effectively delaying the neural signal in that eye

by  $\Delta T_R$ . Neurons tuned to zero disparity would decrease responses (or stop responding) to the pendulum because a stimulus at  $\Delta T_R$  and zero disparity would not be optimal (or excitatory) for them. On the other hand, such a stimulus is now excitatory for neurons tuned to near (crossed) disparities and would elicit responses from them (Fig. 6, top middle). Therefore, the perceived depth would be expected to shift toward the observer. On the other hand, if the filter is placed in front of the left eye, then neurons tuned to far (uncrossed) disparities would be excited (Fig. 6, top right). This would result in a shift of the perceived depth away from the observer. When the pendulum swings to the left, the same principle applies to neurons that prefer leftward motion, except that the direction of shift in the perceived depth reverses (Fig. 6, bottom row). A similar explanation has been proposed based on predictions of a hybrid energy model<sup>18</sup>. Our data provide direct physiological evidence that supports the model.

Another finding reported here is that fine disparities are encoded by neurons tuned to various speeds, whereas coarse disparities are encoded by those tuned to high speeds. Is there an advantage of encoding motion and binocular disparity in this manner? We propose that the relationship between optimal speed and binocular disparity signifies the strategy by which the visual system encodes the three-dimensional structure of a scene. Motion of an observer in a direction other than that along the line of sight creates differences in the motion of retinal images of stationary objects at different depths (motion parallax). This is another cue for depth perception<sup>40</sup>. When the moving observer fixates a point, images of the stationary objects farther away from the fixation point move faster than those of objects closer to the fixation point. In other words, the speed of image motion and binocular disparity are closely correlated. Therefore, if neurons were to encode information about the three-dimensional structure of a scene based on both motion and binocular disparity, then the relationship shown in Fig. 5c would be expected.

There is increasing evidence that suggests that motion and depth are processed together in the visual system<sup>41-46</sup>. Our results strongly support this notion. We show here that phenomena such as the Pulfrich effect and dichoptic motion are likely to be the consequences of joint-encoding at the level of single neurons in the striate cortex. We also report that binocular complex cells in the cat's striate cortex encode specific combinations of motion and binocular disparity that are consistent with the relationship predicted from motion parallax. Therefore, the joint-encoding of motion and binocular disparity is presumably a design of visual processing.

The visual system could process motion and binocular disparity separately before combining the results of the computations together to obtain the three-dimensional structure of a scene. Instead, the visual system integrates them in a specific manner. Such a strategy is not only parsimonious and efficient, it is also reliable; it uses information about the correlation between motion and binocular disparity in the image, which is unavailable if the two cues are processed separately.

#### METHODS

**Electrophysiology.** Extracellular recordings were made from single neurons in the striate cortex of anesthetized and paralyzed cats. Details of surgical and experimental procedures and the physiological recording setup are described elsewhere<sup>20,21</sup>. All procedures complied with the National Institute of Health Guide for the Care and Use of Laboratory Animals and were approved by the University of California Animal Care Committee.

After orientation and spatial frequency tuning of neurons were determined with drifting sinusoidal gratings, responses of binocular neurons



to one-dimensional dichoptic white noise stimuli were measured. Each monocular stimulus consisted of 16 bars. The orientation of the bars was set to the optimal for each cell. The luminance of each bar was updated every 40 ms according to a binary m-sequence<sup>19</sup>, and took a value that was either brighter (+18 cd/m<sup>2</sup>) or darker (-18 cd/m<sup>2</sup>) than the luminance of the background (20 cd/m<sup>2</sup>). The visual stimulation lasts about 20 min, during which spike activity is recorded.

**Data analysis.** Spike trains were cross-correlated with the stimulus sequences of two bars at various spatial offsets and time differences to obtain response profiles for two-bar interaction. When the two bars were chosen from the same eye, the cross-correlation yielded a monocular two-bar interaction profile. On the other hand, if the two bars belonged to different eyes, a binocular interaction profile was obtained. Both monocular and binocular profiles were considered at a common optimal correlation delay, which is defined as the delay at which the sum of squared values of all data points in the profile is maximum. The *x*-axis of the profiles was the relative spatial distance between the two bars ( $\Delta X_L$  and  $\Delta X_R$  for the left and right eye interactions, respectively, and *D* for the binocular interaction, which corresponds to binocular disparity). The *y*-axis was the relative temporal distance between the two bars ( $\Delta T$ ). These profiles were analogous to those reported previously<sup>22,38</sup>. However, they differ from those of our previous study<sup>35</sup>; the *y*-axis was  $\Delta T$ , not a correlation delay *T*. Because of an intrinsic property of binary m-sequences, the interaction of a bar with itself at  $\Delta T = 0$  cannot be obtained. Therefore, one data point is missing from each monocular profile (at  $\Delta X_L$  or  $\Delta X_R = 0$  and  $\Delta T = 0$ ).

Each interaction profile was transformed into the spatio-temporal frequency domain by means of a fast Fourier transform and optimal spatial ( $F_s$ ) and temporal ( $F_t$ ) frequencies are obtained at the peak. Then, the optimal speed (*V*) was computed as  $F_t/F_s$ . Although it is not intuitive to think of spatial and temporal frequencies and speed in the binocular domain, these terminologies are used for both monocular and binocular profiles for simplicity. In addition, for each binocular profile, a slice at  $\Delta T = 0$  is taken, and a one-dimensional Gabor function is fit to the sliced profile using a Levenberg–Marquardt method<sup>47</sup>. The optimal binocular disparity (*d*) was estimated as the phase of the best fit Gabor function, expressed in degree visual angle (degree VA) by taking the optimal spatial frequency of the cell into account.

To quantify the degree of tilt in the binocular profile, a tilt directional index (TDI) was computed as  $TDI = (R_p - R_n)/(R_p + R_n)$ , where  $R_p$  and  $R_n$  are response amplitudes at ( $F_s, F_t$ ) and ( $F_s, -F_t$ ), respectively.

#### ACKNOWLEDGEMENTS

We thank G. DeAngelis for helpful comments and suggestions. This work was supported by research and CORE grants from the National Eye Institute (EY-01175 and EY-03176).

RECEIVED 5 SEPTEMBER 2000; ACCEPTED 28 MARCH 2001

1. Braddick, O. A short-range process in apparent motion. *Vision Res.* **14**, 519–527 (1974).
2. Wheatstone, C. Contributions to the physiology of vision: I. On some remarkable and hitherto unobserved phenomena of binocular vision. *Phil. Trans. Royal Soc. Lond.* **128**, 371–394 (1838).
3. Pulfrich, C. Die Stereoskopie im Dienste der isochromem und heterochromem Photometrie. *Naturwissenschaften* **10**, 553–564 (1922).
4. Julesz, B. & White, B. Short term visual memory and the Pulfrich phenomenon. *Nature* **222**, 639–641 (1969).
5. Morgan, M. J. & Thompson, P. Apparent motion and the Pulfrich effect. *Perception* **4**, 3–18 (1975).
6. Burr, D. C. & Ross, J. How does binocular delay give information about depth? *Vision Res.* **19**, 523–532 (1979).
7. Cynader, M., Gardner, J. & Douglas, R. in *Frontiers in Visual Science* (Cool, S. J. & Smith, E. L., eds.) 373–386 (Springer, Berlin, 1978).
8. Carney, T., Paradiso, M. A. & Freeman, R. D. A physiological correlate of the Pulfrich effect in cortical neurons of the cat. *Vision Res.* **29**, 155–165 (1989).
9. Lit, A. The magnitude of the Pulfrich stereophenomenon as a function of binocular differences in intensity at various levels of illumination. *Am. J. Psych.* **62**, 159–181 (1949).
10. Wilson, J. A. & Anstis, S. M. Visual delay as a function of luminance. *Am. J. Psych.* **82**, 350–358 (1969).

11. Nickalls, R. W. The rotating Pulfrich effect, and a new method of determining visual latency differences. *Vision Res.* **26**, 367–372 (1986).
12. Howard, I. P. & Rogers, B. J. *Binocular Vision and Stereopsis* (Oxford Univ. Press, New York, 1995).
13. Tyler, C. W. Stereopsis in dynamic visual noise. *Nature* **250**, 781–782 (1974).
14. Falk, D. S. Dynamic visual noise and the stereophenomenon: interocular time delays, depth and coherent velocities. *Percept. Psychophys.* **28**, 19–27 (1980).
15. Lee, D. N. A stroboscopic stereophenomenon. *Vision Res.* **10**, 587–593 (1970).
16. Morgan, M. J. Perception of continuity in stroboscopic motion: a temporal frequency analysis. *Vision Res.* **19**, 491–500 (1979).
17. Ross, J. Stereopsis by binocular delay. *Nature* **248**, 363–364 (1974).
18. Qian, N. & Andersen, R. A. A physiological model for motion-stereo integration and a unified explanation of Pulfrich-like phenomena. *Vision Res.* **37**, 1683–1698 (1997).
19. Sutter, E. E. in *Nonlinear Vision: Determination of Neural Receptive Fields* (Pinter, R. B. & Nabet, B., eds.) 171–220 (CRC, Boca Raton, 1992).
20. Anzai, A., Ohzawa, I. & Freeman, R. D. Neural mechanisms for processing binocular information I. Simple cells. *J. Neurophysiol.* **82**, 891–908 (1999).
21. Anzai, A., Ohzawa, I. & Freeman, R. D. Neural mechanisms for processing binocular information II. Complex cells. *J. Neurophysiol.* **82**, 909–924 (1999).
22. Emerson, R. C., Citron, M. C., Vaughn, W. J. & Klein, S. A. Nonlinear directionally selective subunits in complex cells of cat striate cortex. *J. Neurophysiol.* **58**, 33–65 (1987).
23. DeAngelis, G. C., Ohzawa, I. & Freeman, R. D. Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. I. General characteristics and postnatal development. *J. Neurophysiol.* **69**, 1091–1117 (1993).
24. DeAngelis, G. C., Ghose, G. M., Ohzawa, I. & Freeman, R. D. Functional micro-organization of primary visual cortex: receptive field analysis of nearby neurons. *J. Neurosci.* **19**, 4046–4064 (1999).
25. O. Braddick & A. Adlard in *Visual Psychophysics and Physiology* (Armington, J. C., Krauskopf, J. & Wooten, B. R., eds.) 417–426 (Academic, New York, 1978).
26. Georgeson, M. A. & Shackleton, T. M. Monocular motion sensing, binocular motion perception. *Vision Res.* **29**, 1511–1523 (1989).
27. Carney, T. & Shadlen, M. N. Dichoptic activation of the early motion system. *Vision Res.* **33**, 1977–1995 (1993).
28. Carney, T. & Shadlen, M. N. Binocularity of early motion mechanisms: comments on Georgeson and Shackleton. *Vision Res.* **32**, 187–191 (1992).
29. Georgeson, M. A. & Shackleton, T. M. No evidence for dichoptic motion sensing: a reply to Carney and Shadlen. *Vision Res.* **32**, 193–198 (1992).
30. Morgan, M. J. & Castet, E. Stereoscopic depth perception at high velocities. *Nature* **378**, 380–383 (1995).
31. Ziegler, L. R. & Roy, J.-P. Large scale stereopsis and optic flow: depth enhanced by speed and opponent-motion. *Vision Res.* **38**, 1199–1209 (1998).
32. Adelson, E. H. & Bergen, J. R. Spatio-temporal energy models for the perception of motion. *J. Opt. Soc. Am. A2*, 284–299 (1985).
33. Emerson, R. C., Bergen, J. R. & Adelson, R. H. Directionally selective complex cells and the computation of motion energy in cat visual cortex. *Vision Res.* **32**, 203–218 (1992).
34. Ohzawa, I., DeAngelis, G. C. & Freeman, R. D. Stereoscopic depth discrimination in the visual cortex: neurons ideally suited as disparity detectors. *Science* **249**, 1037–1041 (1990).
35. Ohzawa, I., DeAngelis, G. C. & Freeman, R. D. Encoding of binocular disparity by complex cells in the cat's visual cortex. *J. Neurophysiol.* **77**, 2879–2909 (1997).
36. Movshon, J. A., Thompson, I. D. & Tolhurst, D. J. Receptive field organization of complex cells in the cat's striate cortex. *J. Physiol. (Lond.)* **283**, 79–99 (1978).
37. Szulborski, R. G. & Palmer, L. A. The two-dimensional spatial structure of nonlinear subunits in the receptive fields of complex cells. *Vision Res.* **30**, 249–254 (1990).
38. Gaska, J. P., Jacobson, L. D., Chen, H. W. & Pollen, D. A. Space-time spectra of complex cell filters in the macaque monkey: a comparison of results obtained with pseudowhite noise and grating stimuli. *Vis. Neurosci.* **11**, 805–821 (1994).
39. Ohzawa, I. & Freeman, R. D. The binocular organization of complex cells in the cat's visual cortex. *J. Neurophysiol.* **56**, 243–259 (1986).
40. Rogers, B. J. & Graham, M. E. Motion parallax as an independent cue for depth perception. *Perception* **8**, 125–134 (1979).
41. Poggio, G. F. & Fischer, B. Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. *J. Neurophysiol.* **40**, 1392–1405 (1977).
42. Maunsell, J. H. & Van Essen, D. C. 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 (1983).
43. Felleman, D. J. & Van Essen, D. C. Receptive field properties of neurons in area V3 of macaque monkey extrastriate cortex. *J. Neurophysiol.* **57**, 889–920 (1987).
44. Roy, J.-P., Komatsu, H. & Wurtz, R. H. Disparity sensitivity of neurons in monkey extrastriate area MST. *J. Neurosci.* **12**, 2478–2492 (1992).
45. Bradley, D. C., Qian, N. & Andersen, R. A. Integration of motion and stereopsis in middle temporal cortical area of macaques. *Nature* **373**, 609–611 (1995).
46. DeAngelis, G. C., Cumming, B. G. & Newsome, W. T. Cortical area MT and the perception of stereoscopic depth. *Nature* **394**, 677–680 (1998).
47. Press, W. H., Teukolsky, S. A., Vetterling, W. T. & Flannery, B. P. *Numerical Recipes in C* 2nd edn. (Cambridge Univ. Press, New York, 1992).