

Aus dem Physiologischen Institut  
der Ludwig-Maximilians-Universität München  
Vorstand: Prof. Dr. A. Konnerth

Neurons in striate cortex limit the spatial  
and temporal resolution for  
detecting disparity modulation.

Dissertation  
zum Erwerb des Doktorgrades der Medizin  
an der Medizinischen Fakultät der  
Ludwig-Maximilians-Universität zu München

vorgelegt von  
Hendrikje Nienborg

aus  
Freiburg

Jahr 2005

Mit Genehmigung der Medizinischen Fakultät  
der Universität München

Berichterstatter: Prof. Dr. A. Konnerth

Mitberichterstatter: Prof. Dr. A. Mueller  
Prof. Dr. I. Rentschler

Dekan: Prof. Dr. med. D. Reinhardt

Tag der mündlichen Prüfung: 28. 07. 2005

## Zusammenfassung

Stereopsis ist das Sehen räumlicher Tiefe durch binokulare Disparität. Es ist erstaunlich, wie viel weniger fein wir Änderungen binokularer Disparität über Raum (Tyler, 1974; Prince and Rogers, 1998; Banks et al., 2004a) und Zeit (Norcia and Tyler, 1984) wahrnehmen als räumliche und zeitliche Veränderungen von Helligkeitskontrasten. Um die physiologische Grundlage dieses geringen räumlichen und zeitlichen stereoptischen Auflösungsvermögens zu untersuchen, habe ich die neuronale Antwort disparitätssensibler Zellen im primären visuellen Kortex (V1) von vier wachen Affen gemessen.

Zur Untersuchung der physiologischen Grundlagen des räumlichen stereoptischen Auflösungsvermögens, habe ich die dreidimensionale Struktur von 55 rezeptiven Feldern in V1 mit Zufallspunktstereogrammen charakterisiert. Diese Stimuli variierten die Disparität als sinusoidale Funktion vertikaler Position (dichoptisch werden sie deshalb als Welle wahrgenommen). Wenn die räumliche Wellenfrequenz niedrig war, modulierten die Neuronen ihre Entladungsrate mit der zeitlichen Frequenz des Stimulus. Bei höheren räumlichen Frequenzen nahm diese Modulation ab. Die mittlere Entladungsrate änderte sich nur wenig. Sie entsprach ungefähr der Rate, die durch einen gleichförmigen Stimulus mit mittlerer Disparität der Welle hervorgerufen wurde. Bei 48/55 (91%) der Neuronen war die Modulationsstärke eine Tiefpassfunktion der räumlichen Frequenz. Diese Ergebnisse sprechen dafür, dass die rezeptiven Felder fronto-parallel planar ausgerichtet sind, und dass die Zellen weder disparitätsbasierte Umfeldhemmung zeigen noch für Disparitätsgradienten selektiv sind. Dieses Schema sagt eine Beziehung zwischen der Größe des rezeptiven Feldes und der räumlichen Eckfrequenz voraus. Meine unabhängigen davon erhobenen Messungen der rezeptiven Feldgröße standen mit dieser Vorhersage in Einklang. Alle diese Eigenschaften stimmen mit dem Ohzawa-Freeman-Modell („binocular energy model“) überein, das funktionell einer binokularen Kreuzkorrelation entspricht.

Um die physiologischen Grundlagen des zeitlichen stereoptischen Auflösungsvermögens zu untersuchen, bestimmte ich bei 59 Neuronen die Selektivität für die zeitlichen Frequenzen der Disparitätsvariation. Dafür verwendete ich Zufallspunktstereogramme, deren Disparität sich als Sinusfunktion der Zeit änderte. Die zeitlichen Frequenzen, für die die Zellen bei Variation von Disparität bzw. Helligkeitskontrast selektiv waren, korrelierten nicht miteinander, und die Eckfrequenzen waren durchschnittlich niedriger für Disparitätsvariation. Letztere korrelierten negativ mit der Latenz und Geschwindigkeit der neuronalen Antwort sowie mit der zeitlichen Integrationszeit (berechnet aus der Steigung der Phasen-Frequenzfunktion der neuronalen Antwort). Binokulare Kreuzkorrelation der zeitlich bandpassgefilterten monokularen Antworten kann alle diese Ergebnisse erklären.

Die optimale zeitliche Frequenz für Disparitätsvariation war 2Hz. Das entspricht den Werten für die höchste psychophysikalische Sensitivität, die ich bei vier Versuchspersonen ermittelte (1.5-3Hz). Auch die mittlere räumliche Eckfrequenz, 0.5 cpd ("cycles per degree") ist dem Wert, der in psychophysikalischen Experimenten mit

äquivalenten Stimuli gemessen wurde, sehr ähnlich (Tyler, 1974; Prince and Rogers, 1998; Banks et al., 2004a).

Meine Ergebnisse sprechen dafür, dass das menschliche zeitliche und räumliche Auflösungsvermögen für Disparitätsänderungen von Neuronen in V1 begrenzt wird. Ein und derselbe Mechanismus, nämlich binokulare Kreuzkorrelation zwischen den monokularen Abbildern, erklärt sowohl für den Raum als auch für die Zeit, warum die Auflösung für Disparitätsänderungen so viel niedriger ist als die für Kontraständerungen. Gleichzeitig gelingt mit diesen Ergebnisse ein wichtiger Schritt dahin zu verstehen, wie Neurone das Stereokorrespondenzproblem (Julesz, 1971) lösen.

## Abstract

Stereopsis is the process of seeing depth constructed from binocular disparity. The human ability to perceive modulation of disparity over space (Tyler, 1974; Prince and Rogers, 1998; Banks et al., 2004a) and time (Norcia and Tyler, 1984) is surprisingly poor, compared with the ability to detect spatial and temporal modulation of luminance contrast. In order to examine the physiological basis of this poor spatial and temporal resolution of stereopsis, I quantified responses to disparity modulation in disparity selective V1 neurons from four awake behaving monkeys.

To study the physiological basis of the spatial resolution of stereopsis, I characterized the three-dimensional structure of 55 V1 receptive fields (RF) using random dot stereograms in which disparity varied as a sinusoidal function of vertical position (“corrugations”). At low spatial frequencies, this produced a modulation in neuronal firing at the temporal frequency of the stimulus. As the spatial frequency increased, the modulation reduced. The mean response rate changed little, and was close to that produced by a uniform stimulus at the mean disparity of the corrugation. In 48/55 (91%) of the neurons, the modulation strength was a lowpass function of spatial frequency. These results suggest that the neurons have fronto-parallel planar receptive fields, no disparity-based surround inhibition and no selectivity for disparity gradients. This scheme predicts a relationship between RF size and the high frequency cutoff. Comparison with independent measurements of RF size was compatible with this. All of this behavior closely matches the binocular energy model, which functionally corresponds to cross-correlation: the disparity modulated activity of the binocular neuron measures the correlation between the filtered monocular images.

To examine the physiological basis of the temporal resolution of stereopsis, I measured for 59 neurons the temporal frequency tuning with random dot stereograms in which disparity varied as a sinusoidal function of time. Temporal frequency tuning in response to disparity modulation was not correlated with temporal frequency tuning in response to contrast modulation, and had lower temporal frequency high cutoffs on average. The temporal frequency high cut for disparity modulation was negatively correlated with the response latency, the speed of the response onset and the temporal integration time (slope

of the line relating response phase and temporal frequency). Binocular cross-correlation of the monocular images after bandpass filtering can explain all these results.

Average peak temporal frequency in response to disparity modulation was 2Hz, similar to the values I found in four human observers (1.5-3Hz). The mean cutoff spatial frequency, 0.5 cpd, was similar to equivalent measures of decline in human psychophysical sensitivity for such depth corrugations as a function of frequency (Tyler, 1974; Prince and Rogers, 1998; Banks et al., 2004a).

This suggests that the human temporal and spatial resolution for stereopsis is limited by selectivity of V1 neurons. For both, space and time, the lower resolution for disparity modulation than for contrast modulation can be explained by a single mechanism, binocular cross-correlation of the monocular images. The findings also represent a significant step towards understanding the process by which neurons solve the stereo correspondence problem (Julesz, 1971).

## Acknowledgements / Statement of Authorship

In first place, I would like to thank Dr. B.G. Cumming for his excellent supervision throughout. He was an extremely approachable supervisor and extraordinarily reliable to respond promptly to any conceivable question, problem or comment. What was also very important to me was his continuous enthusiasm for the work. I also particularly thank Prof. A.J. Parker and Prof. A. Konnerth for their generous support.

To my family and Stephan.

Some of the data on contrast modulation in Chapter 5 were recorded by Dr. B.G. Cumming. I collected all the rest of the data. I did all of the data-analysis in this thesis using analyzing programs that I wrote myself. I ran the simulations in Chapter 4 on programs written by Dr. H. Bridge. Dr. J.C.A. Read helped me with the mathematics in Appendix B and C. A small number of data has been used previously for a 4 months MSc project. These correspond to 1 datapoint (out of 55) in Figure 3.6 and 4 datapoints (out of 81) in Figure 3.9. I wrote this thesis myself.

# Contents

1	Introduction.....	1
1.1	Binocular disparity.....	1
1.2	Neurophysiological studies of disparity selectivity.....	4
1.2.1	Primary visual cortex.....	4
1.2.2	Selectivity for binocular disparity in V1.....	7
1.2.3	The energy model for binocular disparity.....	8
1.2.4	Selectivity for complex disparity features.....	10
1.2.5	Temporal resolution for disparity modulation.....	11
1.3	Spatial resolution for disparity modulation in human psychophysics.....	11
1.4	Temporal resolution for disparity modulation in human psychophysics.....	12
1.5	Aims of this thesis.....	13
2	General methods.....	14
2.1	Animal training and surgical procedures.....	14
2.2	Stimulus display for electrophysiology.....	14
2.3	Eye movement recordings.....	15
2.4	Unit recording.....	15
2.5	Off-line isolation of single units.....	15
3	Three-dimensional receptive field structure of V1 neurons and spatial acuity for detecting disparity modulation	
3.1	Introduction.....	17
3.2	Methods.....	18
3.2.1	Measurements of disparity tuning functions and responses to depth modulating corrugations.....	18
3.2.2	Data analysis.....	20
3.2.3	Relative modulation.....	21
3.2.4	Corrugation frequency tuning.....	22
3.2.5	Quantitative measurements of receptive field size.....	23
3.2.6	Spatial frequency tuning measurements with gratings.....	23
3.2.7	Model.....	23
3.3	Results.....	24
3.3.1	Neuronal modulation in response to sinusoidal corrugations.....	24
3.3.2	Responses to disparity modulating corrugations of different frequencies.....	27
3.3.3	Determinants of modulation strength.....	38

3.3.4	Comparisons of physiological data with predictions from the model	42
3.3.5	Comparisons of physiological data with results from human psychophysics	44
<b>4</b>	<b>Temporal frequency tuning for disparity modulation</b>	<b>45</b>
4.1	Introduction	45
4.2	Methods	46
4.2.1	Temporal frequency tuning for disparity modulation	46
4.2.2	Temporal frequency tuning for contrast	47
4.2.3	Onset transient	48
4.2.4	Temporal integration time	49
4.2.5	Psychophysical experiments	49
4.2.5	Gaussian fits	50
4.3	Results	51
4.3.1	Temporal frequency tuning in response to disparity modulation	51
4.3.2	Further examination of the temporal properties of the neurons: response onset and -phase	64
4.3.3	No specialization for signaling motion in depth	69
<b>5</b>	<b>Discussion</b>	<b>70</b>
5.1	Summary of the main results	70
5.2	Three-dimensional receptive field structure of V1 neurons and spatial acuity for detecting disparity modulation	71
5.2.1	Future work	75
5.3	Temporal frequency tuning for disparity modulation	76
5.3.1	Future work	78
5.4	An example for the analogy between space and time in the brain	78
5.5	Encoding disparity: binocular cross-correlation as opposed to feature matching	78
5.6	Conclusion	79
	Appendix A	81
	Appendix B	81
	Appendix C	83
	<b>Bibliography</b>	<b>86</b>
	<b>Publications</b>	<b>108</b>

# Chapter 1

## Introduction

“...a painting, though conducted with the greatest art and finished to the last perfection, both with regard to its contours, its lights, its shadows and its colours, can never show a relieve equal to that of the natural objects, unless these be viewed at a distance and with a single eye.”

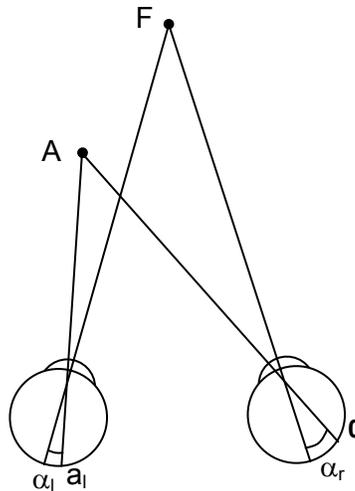
Leonardo da Vinci, *Trattato della Pittura* (cited in Wheatstone, 1838)

### 1.1 Binocular disparity

We usually perceive a visual scene in three-dimensions even though the images on our retinae are two-dimensional. This seemingly geometrical puzzle –seeing in three dimensions when the available visual information is only about two dimensions- is solved because the visual system actively constructs the percept of depth. A number of potential depth cues are available during monocular viewing, such as occlusion of a distant object by a closer one, shading, cues from texture gradients, atmospheric perspective, linear perspective cues, relative size of familiar objects or motion parallax (Gibson, 1950; Rogers and Graham, 1979; Uttal, 1981; Fitzpatrick et al., 1982; Rogers and Graham, 1982; Horn and Brooks, 1989; Anderson and Nakayama, 1994). A particularly potent depth cue requires dichoptic viewing: “binocular disparity”, small positional shifts of matching features in the left and right retinal images relative to the fovea. Stereopsis, the process of seeing depth constructed from binocular disparity was first described by Wheatstone (1838). In this thesis I will study some of the ways in which the

physiological mechanisms for detecting disparity place limits on the psychophysical ability to perceive the disparities.

Since our eyes are located in the front of the head the largest part of the visual field is projected on both retinae. But because of the horizontal separation of the eyes even the overlapping portion of these projections are not identical.



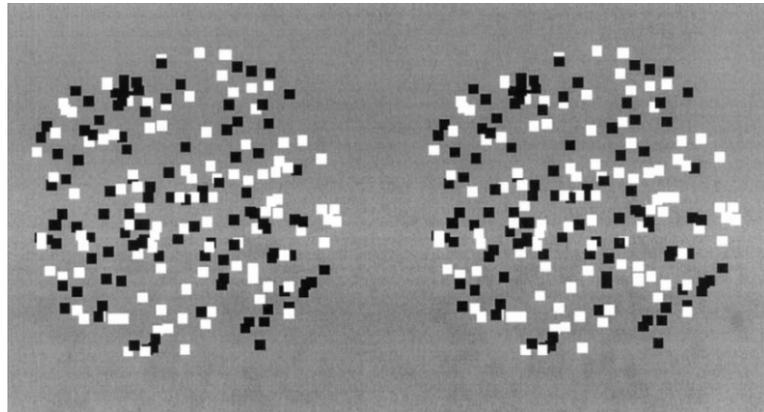
**Figure 1.1** Binocular disparity. The two eyes fixate on F, and the images of F fall on the fovea of both eyes. The positions of the image of point A on the retina are  $a_l$  and  $a_r$ , and their angular distance to the fovea  $\alpha_l$  and  $\alpha_r$ . The difference  $d = \alpha_l - \alpha_r$  is defined as binocular disparity.

Imagine an observer fixating on some point in the visual scene (fixation point F in Figure 1.1). The images of F therefore fall on the fovea of both eyes. And the images of any other point in the visual field (A) fall on the position  $a_l$  and  $a_r$  on the left and right retina relative for the fovea. Taking the fovea as a point of reference, the positions  $a_l$  and  $a_r$  can be expressed as a function of the angles  $\alpha_l$  and  $\alpha_r$ . Binocular disparity is defined as the difference between these two angles. As the separation of the eyes is horizontal the main disparity component during normal viewing is along a horizontal axis (Helmholtz, 1867; Longuet-Higgins, 1982; Read and Cumming, 2004)<sup>1</sup>. Using images of randomly

---

<sup>1</sup> Vertical disparities occur mainly when an object is very close to the observer. They are used to maintain alignment of the eyes (Mayhew and Longuet-Higgins, 1982; Howard and Rogers, 1995) and may play a critical perceptual role under some conditions (Rogers and Bradshaw, 1993; Backus et al., 1999; Berends and Erkelens, 2001; Brenner et al., 2001; Berends et al., 2002). In this thesis I only examined horizontal disparity. If not specified otherwise, I mean horizontal disparity whenever I use the term disparity.

distributed black and white dots (random dot stereograms, RDS), Bela Julesz, (1971) showed that horizontal disparity alone was sufficient to produce the perception of depth.



**Figure 1.2** Random dot stereogram. When the random dot images are fused central circle is perceived as standing out in depth (further away from to the observer).

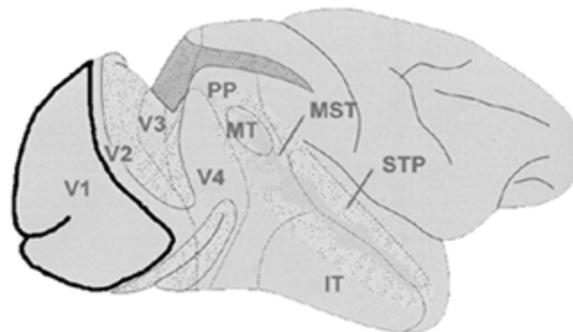
Figure 1.2 shows an example of a random dot stereogram. Both, the left and right panel contain randomly distributed dots. Both panels seem to be identical with all dots being in corresponding positions. But in a circular central region the dots in the right panel are shifted horizontally relative to their counterparts in the left panel. When fusing the two RDS, this gives the impression of a central circle standing out in depth. The information about the circle is not available monocularly. It only is available when the visual system combines information from the two eyes and makes use of the binocular disparity. The ability to perceive depth in these stimuli suggests that the relevant information for stereopsis is extracted early in the visual system, prior to object recognition (Julesz, 1971; Poggio and Poggio, 1984). Random dot stereograms allow one to study the processes relevant for stereopsis independent of the monocular processes. For this thesis all the experiments on binocular disparity and stereopsis were therefore conducted with these kinds of stimuli. Because random dot stereograms contain only two types of features, black and white dots, these stimuli are also apt to study the “stereo correspondence problem” (Julesz, 1971): In a visual scene only one match between features in the left and right eyes’ images gives rise to the appropriate disparity (“correct match”) but there are often other matches possible (“false matches”). In order to perceive depth from images

which contain correct and false matches, the perceptual system therefore needs to solve the correspondence problem and find the appropriate matches.

## 1.2 Neurophysiological studies of disparity selectivity

### 1.2.1 Primary visual cortex

The primary visual cortex (Brodmann area 17, V1) is the largest known visual area in the brain (Zeki, 1990; Tootell et al., 1998), with well defined anatomical boundaries both in the human (e.g. Brodmann, 1909; von Economo and Koskinas, 1925; Stensaas et al., 1974) and in the macaque (reviewed in Felleman and Van Essen, 1991). The macaque striate cortex is largely located within the calcarine sulci and on the operculum of the occipital gyrus, extends rostrally almost to the lunate sulcus and posterolaterally almost to the inferior occipital sulcus (Figure 1.3).



**Figure 1.3** Schematic lateral view of the macaque cortex. The striate cortex (V1, outlined) extends over the majority of the occipital cortex between the lunate sulcus and inferior occipital cortex. After (Albright, 1993).

There is evidence from electrophysiological (Talbot and Marshall, 1941; Daniel and Whitteridge, 1961; Hubel and Wiesel, 1974; Guld and Bertulis, 1976; Van Essen and Maunsell, 1980; Dow et al., 1981), histochemical (Tootell et al., 1982; Tootell et al., 1988b), optical imaging (Grinvald et al., 1986; Blasdel and Salama, 1986; Ts'o et al., 1990) and functional MRI (Dubowitz et al., 2001; Van Essen et al., 2001; Brewer et al., 2002) studies that it contains an entire retinotopically organized map of visual space. V1 is divided between the two hemispheres with each side representing the contralateral half

of visual space. In the macaque, approximately half of the striate cortex on one hemisphere is located on the operculum<sup>2</sup> representing about the central 8 degrees of the contralateral visual field (Blasdel and Campbell, 2001), with the center represented laterally and the visual field inverted around the horizontal axis.

The striate cortex receives its main ascending input from the lateral geniculate nucleus (LGN) (e.g. Hubel and Wiesel, 1972; LeVay and Gilbert, 1976; Hendrickson et al., 1978; Blasdel and Lund, 1983). In the primate, most visual information is funneled through V1 before ultimately reaching the rest of visual cortex (reviewed in Felleman and Van Essen, 1991), which may functionally explain the overwhelming number of connections of V1 with other visual areas (Zeki and Shipp, 1988). For V1, direct connections to (and projections from) have been identified for areas including V2, V3, V3A, V4, V4t, MT, STS, TEO, TE, TO, PO, PIP (Zeki, 1971; Maunsell and van Essen, 1983; Burkhalter et al., 1986; Perkel et al., 1986; Ungerleider and Desimone, 1986a, b; Felleman et al., 1987; Livingstone and Hubel, 1987; Colby et al., 1988; Shipp and Zeki, 1989; Boussaoud et al., 1990; Rockland and Van Hoesen, 1994).

While the input from the ipsi- and contralateral eye remains segregated in different layers in the LGN (Casagrande and Norton, 1991) the majority of the neurons in primary visual cortex receive binocular input (Hubel and Wiesel, 1970). Primary visual cortex is therefore the first stage in the visual pathway where binocular information is combined on one neuron, i.e. thus the first stage where- theoretically- specialization for disparity selectivity is possible.

Visual neurons only respond to stimuli whose projections fall onto a restricted region on the retina. This area is defined as the “receptive field” (Hartline, 1938, 1940) of a neuron. This original rigorous formulation of the receptive field had to be somewhat modified. It is now well established that responses of neurons can be modulated from outside their classical receptive field (Maffei and Fiorentini, 1976; Li and Li, 1994; for review see Allman et al., 1985a), and the non-classical surround of the receptive field itself displays

---

<sup>2</sup> Note that the extension on the operculum in the macaque differs from the human where V1 is mainly restricted to the medial surface on the brain (Wandell, 1999). As a consequence, the representation of the fovea is on the lateral bank of the occipital operculum in the monkey but medial in the human.

variable degrees of temporal and spatial selectivity (e.g. Allman, et al., 1985b; Sillito et al., 1995; Levitt and Lund, 1997; Walker et al., 1999, 2000; Freeman et al., 2001; Cavanaugh et al., 2002; Walker et al., 2002; Bair et al., 2003). These interactions complicate measurements of the receptive field. Nonetheless, the concept of a receptive field remains central for understanding the computations performed by visual neurons. In chapter 3 I will compare predictions of receptive field size obtained for responses to disparity modulating gratings with my independent measurements of receptive field size.

There is a wealth of visual stimulus dimensions for which V1 neurons are specialized, including selectivity for orientation, stimulus size, temporal and spatial frequency, velocity, direction, binocularity, binocular disparity, wavelength (Barlow et al., 1967; Hubel and Wiesel, 1968; Poggio et al., 1975; Schiller et al., 1976a, b, c; Movshon et al., 1978c; Orban et al., 1981; De Valois et al., 1982; Foster et al., 1985; Orban et al., 1986; Tootell et al., 1988a, c-e). The first of these functional specializations characterized for neurons in striate cortex was orientation selectivity (Hubel and Wiesel, 1968). These authors also identified an important distinction between two classes of V1 neurons, simple and complex cells. The responses of simple cells can essentially be described by linear spatio-temporal filtering of contrast (Movshon et al., 1978a; Andrews and Pollen, 1979; Kulikowski and Bishop, 1981; for review see Carandini, 1999). The responses of complex cells are dominated by a non-linear transformation of contrast (Movshon et al., 1978b). These properties have important consequences for the neuronal responses to drifting sinusoidal luminance gratings: the main feature of simple cell responses is modulation at the stimulus frequency ( $f_1$ -component), whereas the response of complex cells is dominated by an unmodulated increase in overall firing ( $f_0$ -component). Relative modulation of the response, defined as the ratio  $f_1$ -component/ $f_0$ -component has been used to classify V1 neurons into simple or complex cells (Movshon et al., 1978a; Skottun et al., 1991). In this thesis I use a similar metric (relative modulation in response to sinusoidal disparity modulation) to measure spatial and temporal frequency tuning to variation in disparity.

## 1.2.2 Selectivity for binocular disparity in V1

Since the first published reports of visual neurons that responded selectively to a particular disparity and not to others in the primary visual cortex of anaesthetized cats (Barlow et al., 1967; Pettigrew et al., 1968) it was speculated that these neurons might represent the basic mechanism underlying depth discrimination. These researchers examined binocular responses to oriented bars influenced by Hubel and Wiesel's reports (Hubel and Wiesel, 1968) that cortical neurons only responded to the appropriate specific stimulus ("trigger feature"), such as bars or edges. After initially being denied in anesthetized monkeys (Hubel and Wiesel, 1970), the existence of disparity selective V1 neurons in response to bars was subsequently confirmed in awake monkeys (Poggio et al., 1977). Poggio and colleagues (Poggio and Talbot, 1981; Poggio et al., 1985; Poggio et al. 1988; Poggio, 1990) found that a large number of the V1 neurons in awake behaving macaques were also selective for binocular disparity when presented with random dot stereograms, i.e. in the absence of any monocular cue. Comparing the range and sharpness of tuning of the cells they recorded to stereoscopic abilities of humans and monkeys, they concluded that V1 neurons were likely to be part of the mechanism leading to depth perception. More recently (Prince et al., 2000) rigorously compared how reliably single neurons discriminate disparity (neurometric thresholds) in alert monkeys with simultaneously measured psychometric thresholds of the animals in a binocular disparity discrimination task. In this study, the signals of a small number of V1 neurons could support the psychophysical precision of stereoacuity. But there is evidence that V1 neurons do not straightforwardly account for the perception of depth from binocular disparity. First, these neurons signal binocular disparity in anticorrelated random dot stereograms (each bright dot is paired with a dark dot and vice versa) (Cumming and Parker, 1997), because these neurons respond to local matches. But anticorrelated RDS do not result in a sensation of depth neither in humans nor monkeys (Cogan et al., 1993; Cumming and Parker, 1997), since there is no "global solution" (Julesz, 1971) to the correspondence problem. [In contrast, neurons with selective responses to correlated but not to anticorrelated RDS have been reported recently in inferior temporal cortex, (Janssen et al., 2003).] Second, it was shown that binocular V1 neurons encoded the local disparity of a grating stimulus (which is ambiguous because of cycle repetitions), not to

the perceived depth of an observer when the stimulus is rendered unambiguous by an aperture (Cumming and Parker, 2000). Finally, there is compelling evidence that stereoscopic depth judgments are based on relative disparities (i.e. the relative disparity between different visual features) rather than absolute disparities (i.e. the disparity relative to the fovea) in humans (Westheimer, 1979; Erkelens and Collewijn, 1985; Kumar and Glaser, 1992) and monkeys (Prince et al., 2000). But the responses of neurons in striate cortex correspond to absolute disparity signals and are not consistent with signaling relative disparity, (Cumming and Parker, 2000).

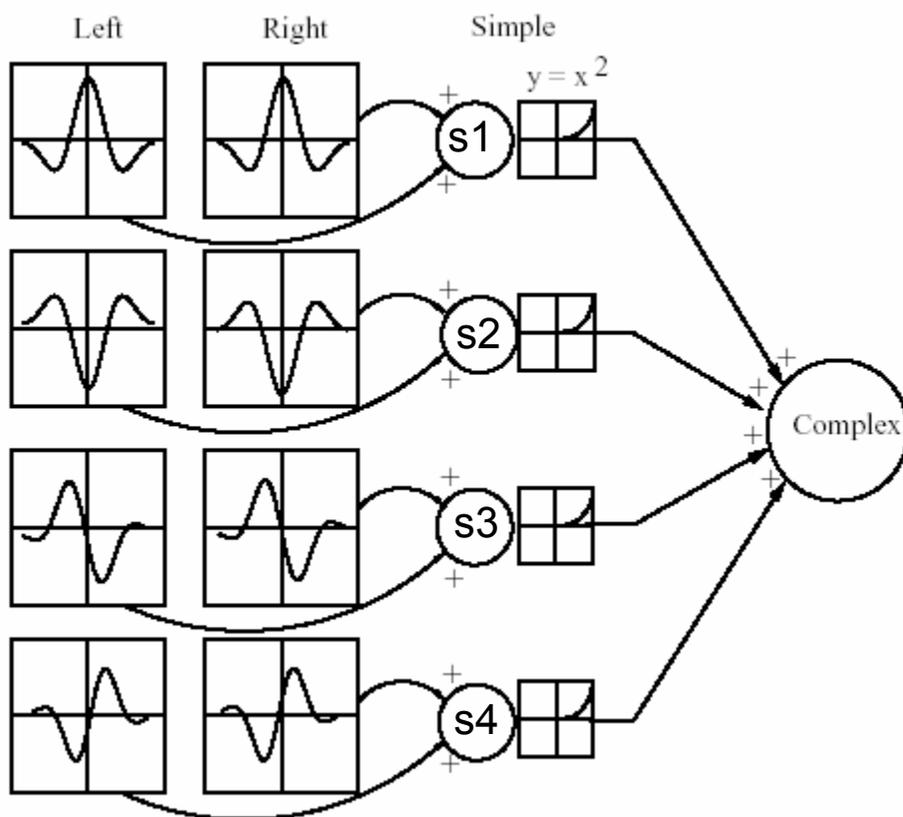
It still seems likely that this early processing of binocular disparity places some limits on stereoscopic performance (Prince et al., 2000; Cumming and DeAngelis, 2001). (In a similar way, the activation of individual photoreceptors is very indirectly linked to our perception of form, but imposes limits on spatial acuity.)

### 1.2.3 The energy model for binocular disparity in V1

Many of the properties of disparity selective neurons in V1 (Ohzawa and Freeman, 1986b; Prince et al., 2002a) can well be described by the binocular energy model (Ohzawa et al., 1990; Fleet et al., 1996; Zhu and Qian, 1996; Ohzawa et al., 1997a, b) (Figure 1.4). This model was initially developed to explain disparity tuning of neurons in area 17 of the cat (Ohzawa, 1990). Figure 1.4 schematically depicts a complex cell selective for zero disparity. It consists of four disparity selective simple-cell sub-units [consistent with quantitative results for simple cells in response to luminance gratings reported by Ohzawa and Freeman, (1986a) for the cat]. The monocular receptive fields of each simple cell are Gabor-shaped spatial filters (DeAngelis et al., 1991). Output from the monocular filters is half-wave rectified (Heeger, 1991, 1993) and squared before converging onto a disparity selective complex cell. All simple cell sub-units are selective for the same disparity (in Figure 1.4, this disparity is zero). When the subunits are in quadrature (90° phase shift of the sine-component of the Gabor-filter between s1 and s3 or s2 and s4 respectively) the complex cell responds independently of stimulus phase. Due to the rectification it is nonetheless selective for disparity<sup>3</sup>.

---

<sup>3</sup> Note that as a first approximation the energy model provides an excellent description of the neuronal data. Nonetheless, a number of quantitative features have been identified (reviewed in Cumming and DeAngelis,



**Figure 1.4** The binocular energy model for disparity. Modified from (Cumming and DeAngelis, 2001) after (Ohzawa et al., 1990). The complex cell is tuned to zero disparity. It is constructed from four binocular simple cells tuned to zero disparity which are in pairs of quadrature phase {s1,s3}, {s2,s4}.

It can be shown that the computation performed by the energy model is equivalent to a cross-correlation of monocular images (Adelson and Bergen, 1985; Fleet et al., 1996; Qian and Zhu, 1997; Anzai et al., 1999a): assume L and R are the output of the left and right monocular receptive fields, respectively of one simple cell subunit. The output of the simple cell subunit is then:

$(L+R)^2 = L^2 + R^2 + 2LR$ , where the last factor,  $2LR$ , causes the simple cell to be disparity selective, and it corresponds to cross-correlation between the left and right images. In chapter 3, I will compare model simulations based on a modification of the binocular

---

2001) that are not captured by this original energy model but by a simple extension invoking a threshold non-linearity before binocular combination (Read et al., 2002; Read and Cumming, 2003b).

energy model using two-dimensional receptive fields (Bridge et al., 2001) with my neuronal results. In chapter 4, I will compare my data with a simulation performing a cross-correlation of the monocular images.

### 1.2.4 Selectivity for complex disparity features

Most studies of disparity selectivity in striate cortex including all studies that have used random dot stereograms have only explored responses to planar, fronto-parallel surfaces at different depths (DeAngelis, 2000). Conversely, typical three-dimensional scenes contain many irregular surfaces. There may therefore be some advantage to constructing receptive fields that are selective for surfaces more complex than fronto-parallel patches. There is evidence that some extrastriate cortical areas show selectivity for three-dimensional slant (Shikata et al., 1996; Sakata et al., 1999; Tsutsui et al., 1999; Hinkle and Connor, 2002; Sugihara et al., 2002; Nguyenkim and DeAngelis, 2003; Taira et al., 2000), or for higher order disparity features (Janssen et al., 1999; Janssen et al., 2000a, b; Janssen et al., 2001; Tanaka et al., 2001). Such responses may simply be inherited from complex three-dimensional receptive fields in V1 (if those exist). Alternatively, the striate cortex may contain a piecewise fronto-parallel depth map, from which the responses in extrastriate cortex are constructed. In such a scheme, the size of the patches in V1 would place limits on the spatial acuity of the three-dimensional maps in extrastriate cortex.

I investigated these questions using a stimulus that permitted me to explore the three-dimensional structure of the receptive field. The stimulus was a random dot pattern in which the disparity was a sinusoidal function of position, producing a grating in depth, (“corrugation”). In principle, the linear superposition of a series of such corrugations can be used to generate any three-dimensional surface [Fourier synthesis in the disparity domain- in analogy to its use for studies of luminance-contrast (Enroth-Cugell and Robson, 1966; Campbell and Robson, 1968; Maffei and Fiorentini, 1973; Kulikowski and Bishop, 1981; Kulikowski et al., 1982; Enroth-Cugell and Robson, 1984; Shapley and Lennie, 1985)]. I will present this work in chapter 3.

### 1.2.5 Temporal resolution for disparity modulation

To my knowledge, there are no published reports on the temporal frequency up to which disparity selective neurons –in striate or extrastriate cortex- signal modulation in disparity.

Instead, it was studied whether neurons were selective for disparity changes over time (motion in depth). Poggio and Talbot, (1981) found a small percentage of macaque V1 neurons sensitive to opposite directions of image motion in the two eyes, and concluded that these cells were tuned for motion in depth. Maunsell and Van Essen, (1983) clarified that selectivity for a fixed disparity would have been sufficient to give rise to an apparent tuning for motion in depth in the above study. These authors recorded from middle temporal neurons of awake macaque monkeys and did not observe selectivity for motion in depth that could not be explained in terms of selectivity for a fixed disparity.

By examining the responses of disparity selective V1 neurons to random dot stereograms whose disparities modulate sinusoidally over time, I will be able to directly address both of these questions. I will do so in chapter 4.

### 1.3 Spatial resolution for disparity modulation in human psychophysics

Tyler, (1973) first studied the spatial acuity for detecting disparity modulation by using vertical line stimuli. If one eye was presented with a straight vertical line and the other with a wavy vertical line, this produced binocularly the percept of a line curved sinusoidally in depth. Tyler found that human subjects could detect depth modulations of the line up to about 3 cycles per degree (cpd). But when the wavy line was viewed monocularly, the undulations could be seen at spatial frequencies above 30 cpd [which agrees with the values for spatial acuity of contrast detection (e.g. Sekuler, 1974)]. Tyler subsequently replicated his findings on the spatial resolution for disparity modulation with random dot stereograms (1974). Presumably because this resolution is surprisingly low when considering the high acuity to detect luminance structure or stereo-acuity, several comparable studies have been conducted (Tyler, 1974; Schumer and Ganz, 1979; Howard and Rogers, 1995; Prince and Rogers, 1998; Bradshaw and Rogers, 1999; Banks

et al., 2004a), with very similar results. Banks et al., (2004a) showed that the poor resolution for disparity modulation cannot be explained by stimulus properties. The stimulus I use to explore the three-dimensional receptive field structure of the V1 neurons is equivalent to the random dot stereograms used in the above psychophysical studies. In chapter 3 I will therefore directly compare my neuronal findings with those from human psychophysics. These data may help to understand some of the limiting factors in the perception of shape from disparity.

## 1.4 Temporal resolution for disparity modulation in human psychophysics

A number of researchers examined how sensitivity for disparity varied as a function of temporal frequency using isolated line (Tyler, 1971; Regan and Beverley, 1973; Beverley and Regan, 1974) or bar stimuli (Regan and Beverley, 1973). Despite considerable differences in the experimental paradigms, these studies roughly agreed that disparity thresholds were lowest between 1-2Hz. For all experiments conducted with line stimuli, the perception of depth failed above 4-6Hz, and in the study using bar stimuli, two transparent bars were seen at different depth above 4Hz. However, in all these experiments the monocular half-images contain a sufficient stimulus for apparent motion. The highest temporal frequency up to which human subjects can see disparity modulation in the absence of monocular motion cues was first investigated by Julesz and Payne using random dot stereograms (Julesz and Payne, 1968). They reported that apparent motion was perceived up to approximately 5Hz and that above frequencies of about 10Hz, two transparent planes were seen. Similarly, Norcia and Tyler (1984) observed 6Hz as the highest frequency for detecting motion in depth. They described that above this frequency, up to about 14Hz “depth pulsations” could be seen. These studies only determined absolute resolution and did not examine the variation of disparity sensitivity as a function of temporal frequency. In chapter 4 I will present disparity sensitivity functions as a function of depth modulation frequency that I measured in four human subjects in order to compare these with the temporal frequency tuning properties I collected for striate neurons.

## 1.5 Aims of this thesis

- 1) To study the three-dimensional receptive field structure of disparity selective neurons of the striate cortex and to investigate the role played by these neurons in the representation of complex three-dimensional surfaces. I will use novel stimuli that allow me to apply the power of Fourier analysis in the disparity domain. I will present this analysis in chapter 3.
- 2) To compare my neuronal results on the three-dimensional receptive field structure with model simulations based on a variant of the energy model for binocular disparity (Bridge et al., 2001). I will include the results of these simulations whenever they are relevant for the interpretation of the neuronal findings and will in addition summarize them at the end of chapter 3.
- 3) To quantify the ability to signal spatial variation of disparity in V1 neurons and to compare these measures with the well known values of human spatial resolution for disparity modulation. I will report this in chapter 3.
- 4) To examine the temporal frequency tuning for disparity modulation of disparity selective neurons in the striate cortex. This is the first study of temporal frequency tuning for disparity modulation. I will try to understand the properties underlying the temporal frequency tuning for disparity modulation by comparing it with temporal frequency tuning in response to contrast, speed of the response onset and temporal integration time. I will also use a novel approach to investigate whether these neurons are selective for signaling motion in depth. I will present this work in chapter 4.
- 5) To measure psychophysically the disparity sensitivity as a function of disparity modulation frequency in human subjects. In order to understand some of the limiting temporal factors for perceiving changes in a three-dimensional scene from disparity, I will compare these results with the temporal frequency tuning in response to disparity modulation. I will present this comparison in chapter 4.

## Chapter 2

### General methods

#### 2.1 Animal training and surgical procedures

For monkeys Hg and Rb, all procedures were carried out in accordance with the U.K. Home Office regulations on animal experimentation. For monkeys Df and Rf, all procedures complied with the Public Health Service policy on the humane care and use of laboratory animals and all protocols were approved by the Institute Animal Care and Use Committee.

I studied neurons in striate cortex (V1) of one female (Rb) and three male (Df, Hg, Rf) alert monkeys (*Macaca mulatta*). In a surgical procedure under general anesthetic, the animals were implanted with a head fixation post and scleral eye coils in both eyes (Judge et al., 1980).

The monkeys were then trained to maintain fixation on binocular visual stimuli for fluid reward.

#### 2.2 Stimulus display for electrophysiology

The visual stimuli were displayed on two EIZO Flexscan 78 monitors (Ishikawa, Japan; for monkeys Hg and Rb) and EIZO Flexscan F980 monitors (for monkeys Df and Rf), and were viewed at a distance of 89 cm in a Wheatstone stereoscope configuration through two small mirrors. The mirrors had a diameter of 18 mm (for monkeys Hg and Rb) and 30 mm (for monkeys Df and Rf), and were positioned about 1.5cm in front of the

animals' eyes. A Silicon Graphics workstation (Mountain View, California) generated the stimuli at a mean luminance of 37 or 42  $\text{cd.m}^{-2}$ , maximum contrast of 99% and at a frame rate of 72 Hz.

### 2.3 Eye movement recordings

A magnetic scleral search system (C-N-C Engineering) measured the positions of both eyes. Only data from trials in which fixation was maintained (to within 0.4 – 1 degree, depending on monkey) were analyzed.

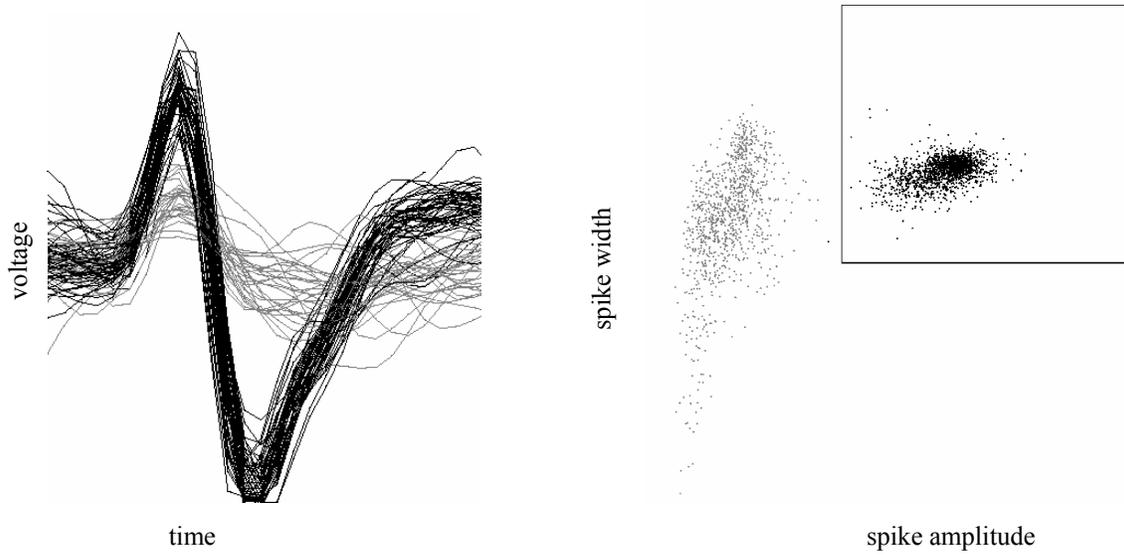
### 2.4 Unit recording

I recorded extracellularly with tungsten in glass microelectrodes (Merrill and Ainsworth, 1972); monkeys Hg and Rb) and glass-coated platinum-iridium electrodes (FHC, Inc.; monkeys Df and Rf), introduced transdurally into the striate cortex each day of recording. The signal was amplified (Bak Electronics, Mount Airy, MD), filtered (200 Hz to 5 kHz), digitized (32 kHz), then stored to disk (using the Datawave discovery system).

The minimum response field of each neuron was initially mapped by hand with a high-contrast bar of approximately preferred orientation. Stimuli were then centered over the minimum response field.

### 2.5 Off-line isolation of single units

Unit isolation was always checked offline. Spike wave forms (voltage changes over time, aligned by their first peak) for one two second trial are shown in the left panel of Figure 2.1. In the right panel two metrics (spike width as a function of spike amplitude) are plotted for each spike form. For the cluster on the right, values of these metrics are similar, and waveforms are considered to originate from a single unit (selected by the black frame). The waveforms corresponding to the values within this window are drawn in black in the left panel.



**Figure 2.1** The left panel depicts spike-waveforms (voltage as a function of time) in one trial superimposed on one another. In the right panel the width (ordinate) is plotted as a function of spike amplitude (abscissa) for each spike. The spikes within the window correspond to spikes with similar values for these metrics and therefore are considered to represent signals from one neuronal unit.

## Chapter 3

# Three-dimensional receptive field structure of V1 neurons and spatial acuity for detecting disparity modulation

### 3.1 Introduction

Although many natural surfaces are slanted or curved in depth, most physiological studies of disparity in V1 have examined neurons only with fronto-parallel surfaces at different depths (DeAngelis, 2000). It has been shown that some extrastriate cortical areas respond selectively to three-dimensional slant (Shikata et al., 1996; Sakata et al., 1999; Tsutsui et al., 1999; Sugihara et al., 2002; Nguyenkim and DeAngelis, 2003), or to the shape and curvature of disparity defined surfaces (Janssen et al., 1999). The question I am asking here is whether such responses are inherited from complex three-dimensional receptive fields in V1 or whether the striate cortex contains a piecewise fronto-parallel depth map, from which the responses in extrastriate cortex are constructed. In such a scheme, the size of the patches in V1 would place limits on the spatial acuity of the three-dimensional maps in extrastriate cortex, and on the cyclopean spatial acuity of the observer.

To investigate this question I used random dot patterns in which disparity varied as a sinusoidal function of vertical position, producing a grating in depth (“corrugation”). In principle, the linear superposition of a series of such corrugations can be used to generate

any three-dimensional surface (Fourier synthesis in the disparity domain). This stimulus therefore allowed me to explore the three-dimensional structure of the receptive field (RF). In order to evaluate the significance of my findings, I also explored the responses of model neurons to disparity corrugations. These were modifications of the energy model (Adelson and Bergen, 1985; Ohzawa et al., 1990; Fleet et al., 1996), using two-dimensional receptive fields (Bridge et al., 2001).

Such corrugation stimuli have also been used to determine the cyclopean spatial acuity of human observers (for review see Rogers and Howard, 1995). I will therefore be able to compare my neuronal results with those of the previous psychophysical studies. These data may help to understand some of the factors limiting the human spatial resolution for detecting disparity modulation. These data are the first to examine neuronal responses to such stimuli, and hence the first time that the power of Fourier analysis has been applied to the study of three-dimensional receptive field structure.

## 3.2 Methods

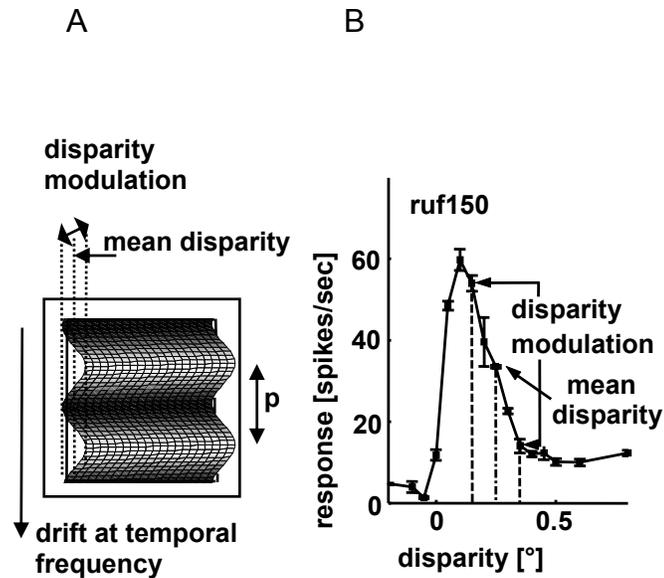
### 3.2.1 Measurements of disparity tuning functions and responses to depth modulating corrugations

Disparity tuning was measured using circular patches of dynamic random dot stereograms (RDS). In order to keep mean luminance constant, the RDS consisted of equal numbers of randomly distributed black and white dots of 99% contrast ( $0.1 \times 0.1^\circ$  size and with an overall density of 50%) on a mid-grey background. A new RDS was presented in each video frame. The RDS was centered on the minimum response field and extended beyond its limits. The disparity of the central region of the RDS varied from trial to trial, while a surrounding annulus ( $0.5\text{--}2^\circ$  wide) was kept at zero disparity. The width of the annulus was always greater than the largest disparity used, eliminating monocularly detectable changes in the stimuli and keeping variation of vergence to a minimum. If the neuron exhibited sensitivity to disparity, responses to sinusoidal modulations of disparity (corrugations, Figure 3.1A) were studied. The disparity defined grating was drifted across the receptive field at a temporal frequency of 2 Hz. The

corrugation had a surrounding rectangular frame at zero disparity. The total stimulus size was  $5^\circ \times 5^\circ$  ( $4^\circ$  corrugation +  $1^\circ$  frame, Figure 3.1A), which always extended well beyond the limits of the minimum response field. Thus any contextual modulation by disparities within  $2^\circ$  of the RF center would be revealed. The corrugation was always oriented horizontally (disparity was a sinusoidal function of vertical position) so that horizontal disparities did not introduce any monocular texture gradients. In a vertically oriented corrugation, the disparity variation leads to changes in dot density, producing monocular images that contain vertical “stripes” defined by regions of high dot density. The only orientation that completely abolishes these monocular changes is horizontal. The disparity modulated between the preferred and the null disparity estimated from the disparity-tuning curve (Figure 3.1B).

The spatial frequency of the corrugation ( $1/\text{spatial period}$ , Figure 3.1A) will be referred to as “corrugation frequency”. It was varied from 0.06 to 4 cycles per degree (cpd) in one octave increments, presented in pseudo-random order. At all corrugation frequencies, the phase of the corrugation was chosen in such a way that the central vertical position of the corrugation was a peak (i.e. phase of  $0^\circ$ ) at the beginning of each 2-second trial. At the lowest corrugation frequency used, the full width of the stimulus was only one quarter of a cycle, so the stimulus over the RF was a nearly uniform disparity with a sinusoidal modulation over time. Each stimulus condition was presented a minimum number of four times (maximum 18 times, mean 8.7).

I interleaved an otherwise identical planar stimulus with a constant disparity equal to the mean disparity of the corrugation. Because there was no temporal variation in the disparity of this stimulus, it allowed me to estimate the extent of temporal modulation in the neuronal firing which was unrelated to the disparity modulation of the stimulus.



**Figure 3.1:** The experimental procedure. A) Schematic representation of the stimulus (corrugation). The mean disparity and the range of disparity modulation were chosen for each cell individually according to the values obtained from the disparity tuning curve (B) of the cell. Corrugation frequency ( $1/p$ ,  $p$ =spatial period) was varied systematically in random order between trials. The planar control stimulus was presented at a disparity equal to the mean disparity of the corrugation. B) An example of the disparity tuning curve for cell ruf150. The arrows indicate the mean disparity of the corrugation and the range for which the depth modulation of the corrugation was chosen for this cell.

### 3.2.2 Data analysis

For most neuronal data, variance increases approximately proportional to mean spike count, (Dean, 1981). Using the square root of the spike counts approximately stabilizes

the variance (Cumming and Parker, 2000; Bridge and Cumming, 2001; Prince et al., 2002b). For this reason, I always used square roots of counts for this analysis or when fitting Gaussians to these data.

All the neurons included in the analysis showed significant modulation of firing rate with changes in disparity (in RDS with uniform disparity) on a one-way ANOVA ( $p < 0.05$ ). An additional selection criterion was that the mean response at the preferred disparity exceeded 15 spikes/sec. The strength of disparity tuning was assessed with the disparity discrimination index (*DDI*, Prince et al., 2002b):

$$DDI = \frac{R_{max} - R_{min}}{R_{max} - R_{min} + 2RMS_{error}} ,$$

where  $R_{max}$  and  $R_{min}$  are the highest and lowest  $\sqrt{rates}$  respectively on the tuning curves, and  $RMS_{error}$  is the square root of the residual variance around the means of  $\sqrt{rates}$  across all disparities. The *DDI* compares the difference in firing to the preferred and null disparity with firing variation. It varies over values between 0 and 1, and can be understood as a metric for the neuron's reliability to signal disparity. If the neuronal firing is strongly modulated by disparity compared to the variability of the response, the *DDI* yields values close to 1. If the neuronal response is not modulated by disparity, any differences in firing to different disparities will be due to noise in the response and the *DDI* will be close to 0.

### 3.2.3 Relative modulation

The corrugation stimuli produced a sinusoidal modulation of the disparity, between the null and preferred values, at the center of the RF. The main manifestation of disparity selectivity in response to these stimuli was therefore a periodic modulation in firing rate at the drift frequency. I quantified this by measuring the *relative modulation* (RM, Movshon et al. 1978). This is the ratio of the amplitude of the modulation at the fundamental frequency ( $f_1$ , the temporal frequency of the drifting corrugation), and mean

response rate ( $f_0$ ). The value for  $f_1$  was obtained by averaging the responses for all trials, then calculating the modulation of this mean response at the frequency  $f_1$ . The same calculation, applied to responses to the planar stimulus, served as a control for the extent of modulation that was not attributable to the corrugation. In order to eliminate artifactual modulation related to response-latency and the onset transient, the first stimulus cycle (500ms) of each trial was discarded.

A resampling (1000 cycles) and bootstrapping method was used (Davison and Hinkley, 1997) to determine confidence intervals. Relative modulation in response to at least one corrugation of the lowest 5 corrugation frequencies had to be significantly higher (on the 5% level, corrected for multiple comparisons) than in response to the planar stimulus, before further analysis of the response modulation was undertaken. This criterion did not reject any neurons that responded only to higher corrugation frequencies - all neurons that showed significant modulation at 2 or 4 cpd showed greater modulation at lower frequencies.

### 3.2.4 Corrugation frequency tuning

For most cells I found that the extent of modulation (RM) decreased as the corrugation frequency of the disparity corrugations was increased (see Results). In order to describe this quantitatively, Gaussian functions were fit to RM as a function of corrugation frequency, using a nonlinear optimization algorithm based on the simplex-search method (Lagarias et al., 1998). Since these were purely descriptive functions, they were allowed to be a Gaussian function of linear or log frequency, whichever minimized the least-square deviations. The Gaussian functions had four free parameters (mean, amplitude, standard-deviation and baseline), which except for the mean were constrained to values  $\geq 0$ .

Gaussian fits were considered adequate if they explained more than 75% of the variance within the data. The high cutoff frequency was defined as the corrugation frequency at which the Gaussian dropped to 2/3 of its peak. (In a few cases, where Gaussians fits had a

mean  $<0$ , it was defined as the frequency at which the Gaussian dropped to 2/3 of the response at 0 cpd)

### 3.2.5 Quantitative measurements of receptive field size

To quantify receptive field size, I measured responses to a thin stationary rectangle ( $0.24 \times 5.99^\circ$ ) of dynamic random dots (each dot usually  $0.08 \times 0.08^\circ$ ) over a range of locations. Receptive field size was measured monocularly for each eye and binocularly, in which case the preferred disparity of the cell was added to the display of the rectangle. In some cases only binocular measures at the preferred disparity were obtained because the cell was lost prior to collecting monocular data or monocular responses were too weak to be reliable. Since I always used horizontally oriented corrugations, response modulation to corrugations provided me with information about the receptive field structure only along a vertical axis. I therefore fit Gaussian functions to the neuronal response rate as a function of the vertical position of the stimulus.

### 3.2.6 Spatial frequency tuning measurements with gratings

The luminance structure of each receptive field was examined with drifting sinusoidal luminance gratings at the preferred orientation and temporal frequency. Spatial frequencies between 0.12 cpd and 16 cpd were presented a minimum of four times each. A Gaussian was fit to the mean response rate as a function of spatial frequency (in linear or logarithmic units, whichever gave the better fit).

### 3.2.7 Model

In order to assist the interpretation of my results, I also examined the responses of model V1 neurons to these stimuli. All simulations were run in Matlab 6 (MathWorks) on either a Pentium PC running Linux or a Silicon Graphics Octane.

The model I used (Bridge et al., 2001) was a version of the binocular energy model (Ohzawa et al., 1990). Two-dimensional Gabor functions were used to simulate receptive fields for each eye. Since the left and right subunits were generated separately,

interocular differences in receptive field structure could be produced. Linear summation of the responses (inner product of the image and the RF) from each eye, followed by halfwave rectification then squaring, generated a binocular simple cell. Summation of four binocular simple cells (in quadrature phase pairs) produced a binocular complex cell.

To create the corrugated RDS, a random number generator was used to determine the x and y positions of the 400 dots, each subtending  $0.1^\circ$  of visual angle. Disparity was added by horizontally displacing dots in opposite directions in the left and right eye stimuli. 500 different RDS patterns were presented to the cells and the responses to each pattern were summed. Relative modulation was calculated as described for the physiological data analysis.

### 3.3 Results

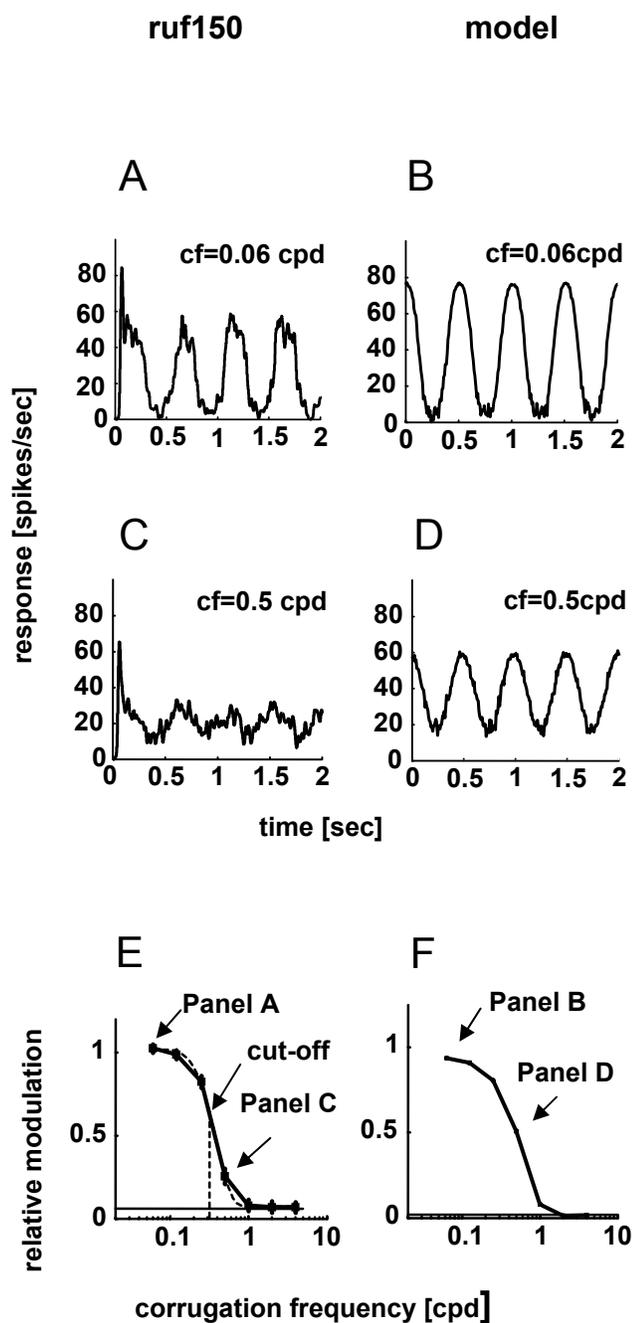
I recorded data from 161 isolated single units in four animals, 62 from monkey Hg, 51 from monkey Df, 36 from monkey Rf and 12 from monkey Rb. Of 106 cells exhibiting significant disparity tuning, 25 failed to yield adequate data for examining responses to depth corrugations (some cells were lost before collecting sufficient data, for others the stimulus was not properly matched to the disparity selective range; one cell was excluded because it showed strong intrinsic modulation of firing close to the stimulus frequency). Of the remaining 81 neurons, 55 showed significant modulation in firing rate when tested with disparity corrugations.

#### 3.3.1 Neuronal modulation in response to sinusoidal corrugations

During presentation of these stimuli, the disparity at any point of the receptive field modulates sinusoidally over time. The temporal frequency of this modulation (2 Hz) is the same throughout the RF. If the neuron's response were a linear function of the disparities of the dots covering the RF, this should result in a sinusoidal modulation of firing rate with the same temporal frequency as the stimulus. This was exactly the pattern I observed (Figure 3.2A). For quantitative analysis of this behavior I calculated the *relative modulation* (RM) – the amplitude of the modulation at the stimulus temporal

frequency ( $f1$ ) divided by the mean firing rate ( $f0$ ). This measure is similar to the F1:F0 ratio measured in response to sinusoidal luminance gratings, which is used to differentiate simple and complex cells (Skottun et al., 1991). An important difference arises because most V1 neurons show little response to a field of uniform luminance, while most neurons in this study showed a substantial response to an RDS with a uniform disparity (whose value was equal to the mean of the disparities covered by my corrugations, see Figure 3.1). This has two significant consequences. First, the value of RM is rarely much larger than 1 (i.e. the neurons are not generally silenced at the null disparity). Second, fluctuations in firing rate in the absence of any changes in disparity mean that the value of RM does not necessarily fall to zero even when there is no disparity modulation. In order to control for this I also measured the response to a corrugation whose depth modulation was 0, i.e. a planar stimulus with a constant disparity over space and time, interleaved with the modulating stimuli.

Because the corrugations were always horizontally oriented, the monocular images defining these stimuli were all indistinguishable. In preliminary experiments responses to other orientations were explored, and in these cases it was necessary to examine responses to the monocular images, since these contain variations in dot density. It was found that in some cases these density variations alone (in monocular images) were sufficient to modulate the neuronal discharge rate. For this reason the rest of this study was limited to horizontally oriented corrugations.



**Figure 3.2:** Responses of a typical neuron (left column, A, C, E) and the energy model (right column, B, D, F) to sinusoidal corrugations in depth. A and B show spike density functions for the response to a corrugation frequency of 0.06 cpd. C and D show responses to a corrugation frequency of 0.5 cpd. The temporal frequency was always 2 Hz.

In panel E and F, relative modulation is plotted as a function of the corrugation frequency. The arrows indicate the relative modulation calculated for the data shown in panels A, B, C, D, and the high corrugation frequency cutoff (0.32 cpd) obtained for cell ruf150 (the corrugation frequency at which the modulation falls to 2/3 of the peak).

### 3.3.2 Responses to disparity modulating corrugations of different frequencies

For the lowest corrugation frequency used (0.06 cpd) there was very little variation in disparity over the extent of the RF of the neurons in this study. This stimulus resembled a traditional planar RDS, in which the disparity of the patch varied over time. The sinusoidal response modulation therefore simply reflects the neuron's selectivity for the disparity of a fronto-parallel planar patch (and the magnitude of the modulation depends on the strength of disparity selectivity). As the modulation amplitude is kept constant and the corrugation frequency increases, spatial variation of disparity within the RF is introduced. At some frequency, this should increase RM in neurons that are selective for any spatial variation in disparity within their RF (just as the structure of a simple cell RF produces the strongest response for intermediate spatial frequency in luminance gratings). If the three-dimensional structure of the RF is planar and selective for the same disparity throughout, the response should depend on the weighted mean of the disparities (the weights being set according to the position of the dots over the RF):

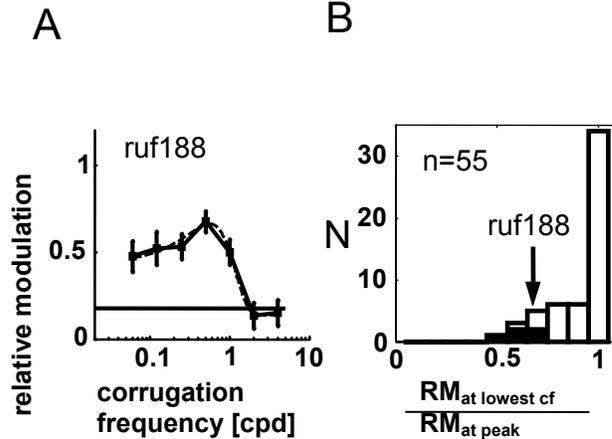
$$R = \frac{1}{n} \sum_{i=1}^n (d(i) \cdot w(i)),$$

where  $R$  denotes the instantaneous neuronal response,  $i$  corresponds to the position in the receptive field,  $d(i)$  to the disparity of the stimulus at this position,  $w(i)$  to the weight of the RF envelope at this position. (This is only true for disparities limited to the linear portion of the response that lies between null and preferred disparities, as in my experiments.) Consider a hypothetical cell with a rectangular RF, i.e. disparity input has equal weights throughout the RF ( $w(i)=\text{constant}$ ). The response then simply depends on the mean disparity, so that when the spatial period of the corrugation equals the RF width, there is no modulation in the mean disparity, and hence no modulation in the response. For a Gaussian RF, modulation decreases as a Gaussian function of increasing

corrugation frequency. Figure 3.2E shows the value of RM as a function of corrugation frequency, for a typical neuron and Figure 3.2F shows the response of the implementation of the energy model. Both show a lowpass function, behaving as if the response at any time is determined by the mean of the disparities over the RF.

A second feature of responses reflecting the mean weighted disparity is that the mean firing rate should not change as a function of corrugation frequency (provided that a full number of corrugation-cycles is presented for each frequency, as I did in my experiments). This was a consistent feature of the neuronal responses - only 9/55 (16%) cells showed a significant change in mean spike rate with corrugation frequency (5% significance on ANOVA). Even these cells showed only modest changes in mean firing rate, and I did not observe a consistent pattern of change in mean firing with respect to corrugation frequency.

I completed tests with corrugations of several different corrugation frequencies for 55 cells that showed significant modulation for at least one frequency. The variation in RM as a function of corrugation frequency was generally well described by Gaussian functions (see Figure 3.2E). All Gaussian fits explained >75% of the variance in the data. These fits were then used to quantify the two main features of interest shown in Figure 3.2E. First, I examined whether all cells had their largest responses at low corrugation frequencies (spatially lowpass). Second, I examined the highest corrugation frequency that produced substantial modulation (high corrugation frequency cutoff, the point at which RM falls to 2/3 of the peak).



**Figure 3.3:** Response attenuation at low frequencies was unusual. Panel A shows relative modulation as a function of corrugation frequency for one neuron (ruf188) with unusually marked attenuation at low frequencies. The dashed line corresponds to the Gaussian curve fitted to the data. The relative modulation for the lowest frequency tested (0.06 cpd) was only 70% of that at the peak of the fit. This ratio, (RM at the lowest frequency tested)/(RM at peak) was used to estimate the extent of low frequency attenuation for the population of 55 neurons (frequency histogram shown in B). Black bars correspond to the ratios of cells with statistically significant attenuation at low frequency.

The fitted Gaussian functions were not constrained to be lowpass – Figure 3.3A shows an example in which the best fit had a peak at intermediate frequencies. I estimated the extent of attenuation in RM at low frequencies by comparing the RM at the lowest frequency tested with the peak value of RM. (Values for this comparison were obtained from the data, not from the fits). The large majority of these ratios were close to 1 ( $> 0.9$  for 40/55 cells, 73%, see Figure 3.3B). Except for one cell, all values were greater than 0.5, and hence would be considered lowpass by the criterion of Hawken et al., (1996).

The apparent reduction in RM at low frequencies for some cells may largely be the result of random fluctuations in response. To evaluate this statistically, I compared the RM at 0.06 cpd with the largest RM observed. This was significant for 5 of the 55 cells ( $p < 0.05$ , two tailed test corrected for multiple comparisons, by resampling). (I also used

nested Gaussian fits to see if adding a non-zero mean produced a significant improvement, with similar results, 7/55 being significant.)

Thus it appears that a small population of neurons does show significant attenuation of their response modulation at low corrugation frequencies. However, even in this selected population the attenuation was modest (the example shown in Figure 3.3A was one of the five significant cases). Thus it does not appear to represent a specialization for processing more complex three-dimensional configurations. The phase of the responses suggested that these were not responses to surface slant: the absolute response phases were close to  $0^\circ$  (response peak when the preferred disparity was in the center of the RF), and there was no systematic shift in phase as a function of corrugation frequency. These features are compatible with a degree of surround suppression produced when the surround was at the same disparity as the center.

The above analyses establish that the vast majority of V1 neurons show a lowpass response to sinusoidal variations in disparity, suggesting fronto-parallel RFs. This is in agreement with previous reports of binocular interaction profiles elongated in the fronto-parallel plane (Anzai et al. 1999, their Figure 1). The results presented here go beyond the previous report in two respects. The examples shown by Anzai et al. display elongation in the fronto-parallel plane, but their analysis did not examine this feature quantitatively or rule out slant-selectivity. The quantitative analysis in the present study suggests that selectivity for a uniform disparity across the RF is a general feature of disparity selective neurons in striate cortex. Further, Anzai and colleagues' use of a reverse-correlation analysis relies on the assumption of a linear neuronal response. They used a sequence of noise stimuli in which there is no single consistent pattern of disparities- whether it be planar or sinusoidal. Thus, a detecting system that is sensitive to a consistent pattern of disparities would be inadequately stimulated by such a noise stimulus. Non-linear responses to combinations of disparities simultaneously present in the RF could therefore not be examined. For example a non-linear interaction between neighboring parts of the RF might generate selectivity for slant independent of mean disparity. Such a response would be revealed by my measures with corrugations, since the stimulus contains

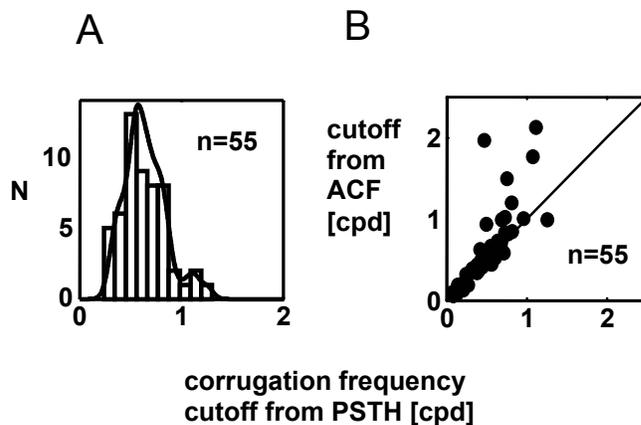
appropriate multiple disparities. The lowpass response to the corrugations therefore allows us to exclude the possibility of non-linear responses to combinations of disparities. It is important to stress that the lowpass response not only indicates linear summation of disparities, but also that the input to this linear summation appears to show the same disparity preference at each point in the RF. Linear summation across an RF with a preferred disparity that changes across the RF would also generate bandpass responses.

Since RM is a lowpass function of corrugation frequency for most neurons, their response can be summarized by the high corrugation frequency cutoff for the RM (see Figure 3.2C). I took as my measure the frequency at which the fitted Gaussian fell to 2/3 of its peak value. Figure 3.4A shows the distribution of this high-cut frequency for 55 cells. It tends to occur at quite low frequencies, with a mean of 0.5 cpd ( $\pm 0.26$  cpd, S.D.), much lower than typical values for luminance gratings in monkey V1 at these eccentricities (Foster et al., 1985). The highest value observed was 1.3 cpd.

However, this analysis is complicated by small eye movements during fixation. Changes in vertical eye position will change the phase of the sinusoid that falls over the RF center, changing the phase of the neuronal response. Significant variation in the phase of the response within and between trials leads to an underestimate of the modulation amplitude. This problem is greatest for high corrugation frequencies, since small movements constitute a larger fraction of the spatial period. This could therefore lead to an underestimate of the high corrugation frequency cutoff. However, the majority of the cutoff frequencies observed were too low to be explained by the observed variation in vertical eye position: the mean of the observed S.D.s was  $0.14^\circ$ ,  $0.19^\circ$ ,  $0.21^\circ$  and  $0.22^\circ$  for monkeys Hg, Rb, Df, and Rf respectively. Since the measured variation in eye position overestimates the true variation (Read and Cumming, 2003a), it seems that eye movements are unlikely to have compromised these measurements in more than a few neurons.

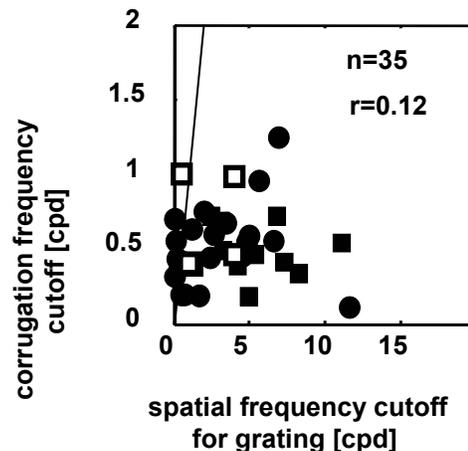
As an additional safeguard against the effect of eye movements, I re-analyzed the data using the autocorrelation function to detect periodic modulation in firing rate,

independent of phase (Schlittgen and Streitberg, 2001). Appendix A shows how with the help of the autocorrelation-function the effect of saccadic eye-movements on relative modulation is reduced. Figure 3.4B compares the value of high corrugation frequency cutoff for the two analysis methods. In all but a few cases, the two methods gave similar answers, confirming that the influence of eye movements on the measurement of RM was generally small. In a few cases, the autocorrelation method produces clearly larger values, suggesting that eye movements may have had a significant effect in these neurons. This leads to a weak, but significant ( $p < 0.01$ , paired t-test) tendency towards higher values for the autocorrelation method (mean value  $0.61 \pm 0.44$  cpd, S.D. by the autocorrelation method).



**Figure 3.4:** High corrugation frequency cutoff. Panel A shows a frequency histogram and superimposes a smoothed density function. It is possible that in some cases, modulation to high frequency corrugations was underestimated because of the effects of fixational eye movements. I therefore re-analyzed all the data using the autocorrelation function of the spike trains to estimate periodic modulation. This is less sensitive to eye movements (see Appendix). Panel B compares the corrugation frequency high cutoffs calculated with this autocorrelation method with the cutoff estimated from the modulation in the peri-stimulus time histogram (PSTH). Despite the generally excellent agreement between the two measures, the cutoff frequency estimated from the autocorrelation functions is substantially larger in a few cases. This leads to a significant difference between the two measures across the population ( $p < 0.01$ ,  $n = 55$ , paired t-test).

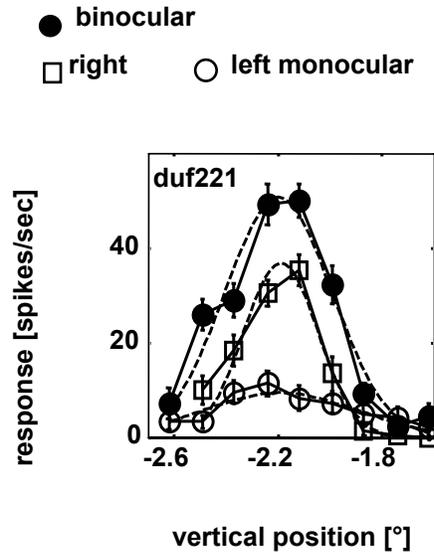
I compared the high frequency cutoff for depth corrugations with the spatial frequency high cutoff for luminance gratings in 35 of the 55 cells. The luminance gratings were presented at the preferred orientation (20 monocular, dominant eye, 15 binocular at the preferred disparity), while the depth corrugations were always horizontal ( $0^\circ$ , disparity variation always along a vertical axis). In order to compare these values therefore, the high-cut frequency for the gratings was multiplied by  $\cos(\text{grating orientation})$ , to estimate the cutoff frequency along the vertical axis. Despite the fact that this shifts the cutoff spatial frequency to lower values for the luminance gratings, these values are almost always greater than the cutoff corrugation frequency for depth modulation (Figure 3.5, note the different scales on abscissa and ordinate). This indicates that the three dimensional structure of the RF is on a coarser scale than the luminance structure. Furthermore, there is no significant correlation between these measures ( $r = 0.12$ , n.s.). Both of these observations follow if the luminance structure of the RF is used to calculate interocular correlation across finite regions of the two retinae (as in the energy model, see simulations below). The three-dimensional structure is determined by the area over which this is estimated, not by the luminance RF structure within it.



**Figure 3.5:** Comparison of high corrugation frequency cutoff values in response to corrugations (ordinate) and high spatial frequency cutoff to sinusoidal luminance gratings (abscissa). The filled symbols show data for 35 neurons, obtained monocularly for the dominant eye (circles,  $n=20$ ) or binocularly (squares,  $n=15$ ).

There is no significant correlation between the two properties of the neurons ( $r=0.12$ ; n.s.,  $n=35$ ). Open squares represent values obtained from the model using four different receptive fields. This illustrates that the model does not require a correlation between corrugation and spatial frequency cutoffs for these two stimuli.

If the neurons were constructed from subunits that are all selective for the same fronto-parallel disparity (as in the energy model, producing a lowpass response) then the high corrugation frequency cutoff for disparity modulation should be determined by the receptive field size. I therefore obtained quantitative estimates of the vertical extent of the receptive field for 21 neurons. These measures used thin strips of random dot texture like those used for the depth corrugations (see Methods). Gaussian functions were fit to the response rate as a function of vertical position, and the S.D. used as a measure of the halfwidth of the minimum response field. Data for one example neuron are shown in Figure 3.6. Two of these 21 neurons showed a degree of bandpass response (suggesting a center-surround organization). Their high frequency cutoff is determined by the size of the center region. [Similarly for retinal ganglion cells, the cutoff spatial frequency is a predictor of center size of the receptive field, e.g. (Linsenmeier et al., 1982).] Since the size of the center of the preferred disparity would have been difficult to measure independently I excluded these two neurons from the comparison.



**Figure 3.6:** Example data illustrating the measurement of receptive field height for cell duf221. Vertical position of the stimulus relative to the fovea is depicted on the abscissa and firing rate on the ordinate. Filled circles represent the measurements in response to a binocular stimulus, open markers in response to a stimulus being presented monocularly. Gaussian functions (dashed lines) are fitted to each of the data-curves.

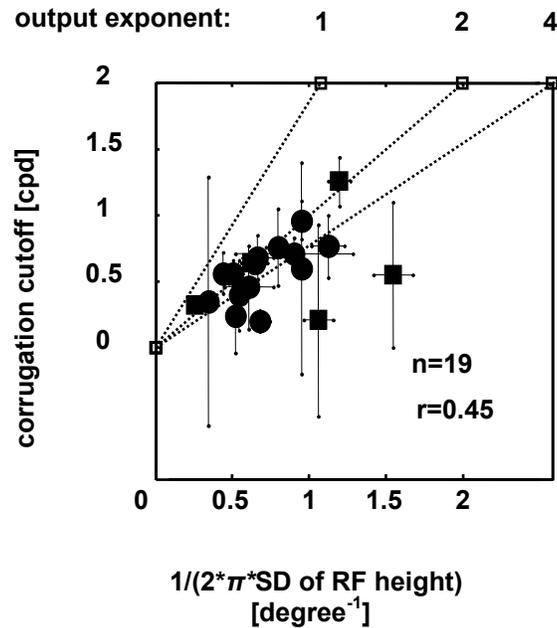
Suppose the response rate of a neuron was given by the sum of its responses to the disparities of individual dots, weighted by the Gaussian envelope of the RF. Under these circumstances, I can use the Fourier transform of the minimum response field to predict the way in which RM varies with corrugation frequency. The Fourier transform of a Gaussian in space with an S.D. of  $\sigma_x$  is a Gaussian in the frequency domain with S.D.

$$\sigma_s = \frac{1}{2\pi\sigma_x}$$

The frequency domain counterpart of the RF envelope is the corrugation frequency tuning width. If RM is a lowpass function of corrugation frequency (as for the vast majority of the neurons and in the energy model), the Gaussian in the frequency domain is centered around zero. In this case frequency tuning halfwidth can be estimated by the corrugation frequency cutoff.

Before examining this relationship for real neurons, I ran simulations with the implementation of the energy model (Bridge et al, 2001), measuring the minimum response field and the high corrugation frequency cutoff for corrugations in the same way as for the neurons. One model parameter that had an important influence on this relationship was the value of the output exponent. (In Appendix B it is shown that for an output exponent of 2, the prediction should correspond to the identity line, which is also found in the simulation.) Although the energy model is usually described with a half-squaring output nonlinearity, this is largely for mathematical convenience. Raising the output to higher powers produces qualitatively similar results, and a number of studies have suggested that the output exponents often are  $>2$  in real neurons (Anzai et al., 1999b; Gardner et al., 1999). For model neurons, larger output exponents lead to smaller estimates of the minimum response field, while the cutoff frequency for depth corrugations was less affected. The relationship between minimum response field, high-frequency cutoff, and output exponent for model neurons is shown by the dotted lines in Figure 3.7.

Superimposed on these model data, Figure 3.7 compares the size of the minimum response field with the high corrugation frequency cutoff for 19 neurons. Only fits that explained  $>70\%$  of the variance were used. The correlation ( $r=0.45$ ,  $p<0.05$ ) is significant and the great majority of neurons (18/19) lie beneath the line obtained from the model-simulations (dotted line at output exponent of 2). The point above the line does not differ significantly from it (by resampling). Cells that fall below the predicted line can readily be explained by assuming an output exponent that is  $> 2$ . Across the population there is a tendency for the high corrugation frequency cutoff to be lower than predicted by the model with an exponent of 2. This may indicate that most neurons have an output nonlinearity that is more expansive than half squaring. Alternatively, it may be that the area over which the disparity integration takes place is somewhat larger than my estimate of the minimum response field.



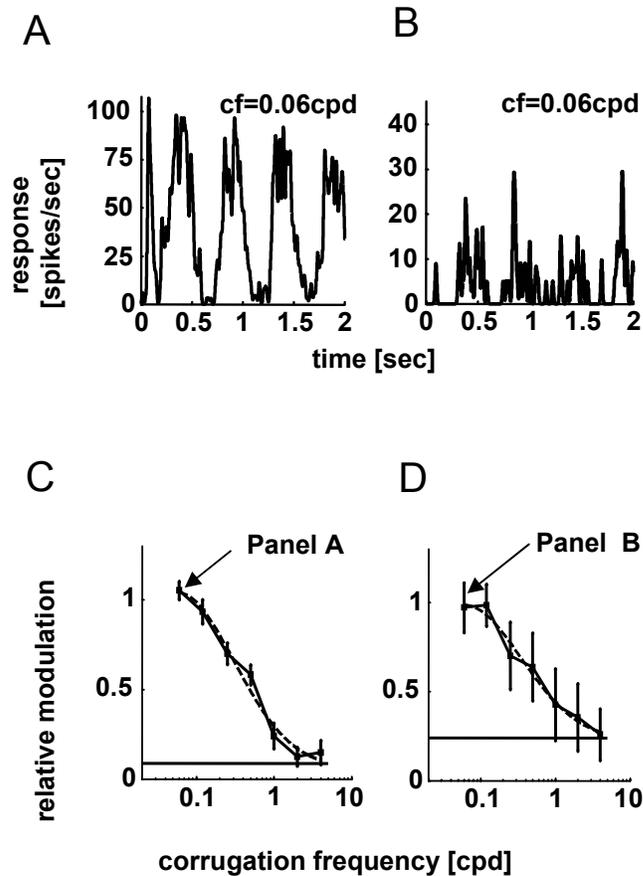
**Figure 3.7:** The relationship between receptive field size (the S.D. of a Gaussian fit to responses like those in Figure 3.6,  $SD_{rf}$ ) and high frequency cutoff for depth corrugations (the S.D. of a lowpass Gaussian fit to data like those shown in Figure 3.2E,  $SD_{sf}$ ). Since an inverse relationship is expected,  $1 / (2\pi * SD_{rf})$  is plotted on the abscissa (i.e. the S.D. of the same Gaussian in the frequency domain). There is a significant correlation, with ( $r=0.45$ ,  $p<0.05$ ,  $n=19$ ). The solid line represents the identity-line. Circles and squares depict cells, for which the receptive field was measured monocularly in the dominant eye ( $n=14$ ) and binocularly ( $n=5$ ), respectively, error bars correspond to standard-errors (by resampling).

This relationship was also examined with the model (dotted lines). The model responses are shown for three different output exponents (1,2 and 4). Note that no neuron lies significantly above the line predicted by the model with an output exponent of 2 (i.e. a half-squaring output nonlinearity). This line also represents the identity line. The output exponent of 1 refers simply to half-wave rectification, a sufficient nonlinearity to generate disparity selectivity.

Although the high corrugation frequency cutoff was significantly correlated with minimum response field size, there was considerable scatter in this relationship. This may explain why I found no clear relationship between the high corrugation frequency cutoff and eccentricity ( $r=0.04$ ,  $n=55$ , n.s. partial correlation with respect to monkey).

### 3.3.3 Determinants of modulation strength

The analysis above has primarily focused on how RM varies with changes in the corrugation frequency of depth modulation. The strength of the modulation was largely ignored. In this context it is important to note that the estimation of RM is more reliable with higher mean firing rates. Figure 3.8 demonstrates this for two neurons firing at a mean spike rate of 46 spikes/sec (hg599, left column), and 8 spikes/sec (hg620, right column) respectively. Both neurons produce similar values of RM. The 2 Hz modulation is however much clearer in the spike density function for the neuron with the higher firing rate, (Figure 3.8A). The responses to the planar control stimulus show no clear modulation in either cell (data not shown), but their RM values differ by a factor of 3 (baseline in Figure 3.8C and 8D: RM=0.08 for cell hg599, RM=0.24 for cell hg620).



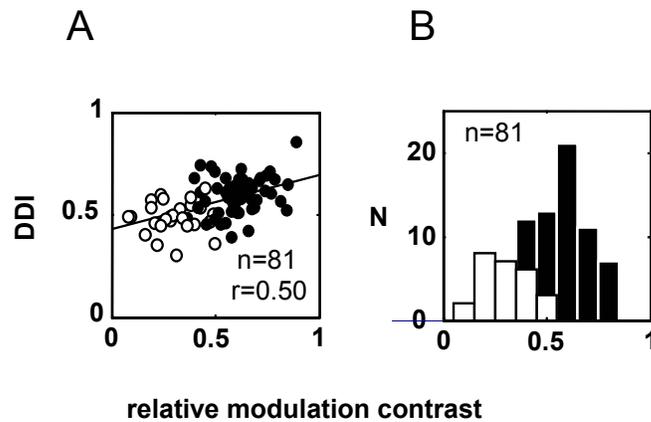
**Figure 3.8:** Relative modulation and mean firing rate. Panels A and B show spike density functions for two neurons, with very different mean firing rates (46 spikes/sec and 8 spikes/sec respectively). The corrugation frequency of the disparity corrugation was 0.06 cpd, the temporal frequency was 2 Hz. Although the values of relative modulation are similar for both neurons (C,D), the modulation is much clearer for the neuron with the higher firing rate (A). Neurons with lower firing rates give less reliable estimates of relative modulation. This is most clearly seen from the error bars in (C, D), which show S.D.s of the resampled populations

Raw values of RM can therefore be a misleading guide as large values can occur at random in weakly modulating cells with low mean firing, or simply because a neuron has a tendency to periodic activity regardless of the stimulus. I therefore quantified the modulation strength with a contrast measure, the *relative modulation contrast* which compares the modulation produced by corrugations with that produced by the planar stimulus with no disparity variation:

$$\text{relative modulation contrast} = \frac{RM_{max} - RM_{planar}}{RM_{max} + RM_{planar}}$$

where  $RM_{max}$  is the peak RM in response to the corrugation and  $RM_{planar}$  is the RM in response to the planar stimulus with no temporal or spatial modulation in disparity.

Note that when neuronal firing is weak, the value of RM for the planar control stimulus is often large, so this measure is not biased towards large values in weakly activated neurons [unlike similar measures on raw spike counts, (Prince et al., 2002b)]. I investigated the relationship between relative modulation contrast and two properties: retinal eccentricity, and disparity selectivity (DDI, Figure 3.9). There is no significant dependence of modulation on retinal eccentricity ( $r=-0.0007$ ,  $n=55$ , n.s., partial correlation with respect to monkeys, data not shown). But there is a highly significant positive correlation with the DDI ( $r=0.50$ ,  $p<10^{-5}$ , Figure 3.9A).



**Figure 3.9:** The relationship between strength of disparity tuning (measured by the DDI) and the strength of modulation induced by the corrugations (measured by relative modulation contrast). Filled symbols represent the neurons for which the relative modulation was significant  $n=55$ , open circles indicate the neurons for which relative modulation was not significant,  $n=26$ . All neurons included in this plot had significant

disparity tuning,  $n=81$ . There is a strong and significant ( $r=0.5$ ,  $p<10^{-5}$ ) correlation. All neurons with a DDI greater than 0.65 showed significant relative modulation.

Neurons without significantly higher modulation in response to the corrugations than in response to the planar stimulus (open circles, Figure 3.9A) have a significantly lower DDI (mean DDI 0.30 versus 0.61 for neurons with significant modulation,  $p<10^{-6}$ ,  $n=81$ , t-test). All neurons with a DDI above 0.65 show significant modulation. At first sight this correlation is not surprising, since both measures are closely dependent on the strength of disparity tuning. But this need not be true. Consider a neuron with a response that depends primarily on the magnitude of surface slant. Such neurons could show weak disparity selectivity for planar RDS, but modulate their activity strongly during presentation of depth corrugations, as these present different values of slant at different points in the cycle.

The correlation between DDI and relative modulation contrast, especially the lack of neurons showing strong relative modulation contrast with weak DDI, is further evidence that there is not a population of neurons in V1 that are selective for slant but only weakly selective for disparity in planar RDS. I also studied responses in 19 neurons that were not disparity selective (ANOVA  $p > 0.05$ ), which are not shown in Figure 3.9. None of these showed significant modulation to the corrugations.

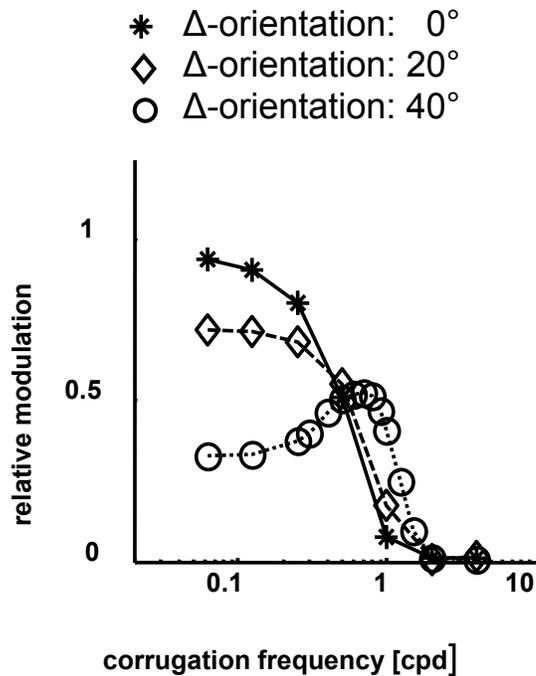
One potentially important source of scatter in the relationship between the strength of disparity tuning and the strength of modulation to depth corrugations may be the temporal properties of the stimulus. All of the data reported here used a stimulus temporal frequency of 2 Hz. This is well below the temporal frequency high cutoff for most neurons when examined with luminance gratings. [All of the neurons reported by Hawken et al., (1996) had temporal frequency high cutoffs  $> 2$  Hz.] However, it is possible that the temporal integration for disparity processing is over greater periods, and that some neurons might have shown stronger modulation had I employed lower temporal frequencies

### 3.3.4 Comparisons of the physiological data with predictions from the model

At several points I have shown responses of model complex cells, using a model similar to the energy-model for disparity tuned complex cells implemented by Ohzawa et al., (1990), but using two-dimensional receptive fields (Bridge et al., 2001). These simulations showed that all of the physiological data are compatible with such a model:

1. The model responded with a sinusoidal modulation of the response at the stimulus temporal frequency (Figure 3.2).
2. The strength of this modulation depended on the relationship between the corrugation frequency of the corrugation and the size of the receptive field, in a similar fashion to that seen in the neuronal data (Figure 3.7).
3. Corrugation frequency selectivity for depth corrugations is determined by RF size. Spatial frequency selectivity for contrast gratings is determined by the structure within the RF. For this reason, spatial frequency tuning for luminance contrast and corrugation frequency for depth corrugations can be manipulated separately in the model (Figure 3.5, open squares). They are also uncorrelated in the physiological data (Figure 3.5).
4. The model shows a lowpass response to depth corrugations. Interestingly, this remains true even when an orientation difference between the eyes is introduced to the model (Figure 3.10). Although it has been suggested that such orientation differences might be a specialization for detecting surface slant (Blakemore et al., 1972; von der Heydt, 1978; Ninio, 1985; Mitchison and McKee, 1990; Cagenello and Rogers, 1993), within the context of the energy model the response remains dominated by the effects of positional disparity [Bridge et al., 2001); their Figure 7], so that the modulation at low frequencies is very similar to that at intermediate frequencies. If very large orientation differences are used (40 degrees, much larger than those found physiologically, (Bridge and Cumming, 2001) a modest increase in response can be seen at intermediate frequencies (Figure 3.10, dotted line). Thus the lack of any bandpass response to disparity corrugations in V1 neurons is not evidence against the existence of interocular differences in preferred

orientation. It is evidence against the existence of neurons primarily sensitive to surface slant.



**Figure 3.10:** The effect of orientation differences in model neurons. Model neurons were constructed with different receptive field orientations in the two eyes. The response of these models to corrugations of variable corrugation frequency was then assessed, using the same protocol as I used for real neurons. Relative modulation (ordinate) obtained from model-simulations is plotted against corrugation frequency (abscissa). Orientation difference between the monocular receptive fields ( $0^\circ$ ,  $20^\circ$  and  $40^\circ$  for solid, dashed and dotted line respectively) was the only parameter changed between the simulations. Note that a decline in relative modulation at low spatial frequencies only becomes apparent for large orientation differences ( $40^\circ$ , open circles).

### 3.3.5 Comparisons of the physiological data with results from human psychophysics

Comparing the physiological data here with psychophysical responses at similar eccentricities suggests that spatial integration of these V1 neurons is what limits performance. The neuronal data show a reduction to 66% of maximal sensitivity at a mean frequency of 0.5 cpd, at a mean eccentricity of 3.7°. If psychophysical judgments depended on these responses, a corrugation frequency of 0.5 cpd should produce thresholds 50% higher than the optimum. The two human studies that examined thresholds at these eccentricities (Prince and Rogers, 1998; Banks et al., 2004a) found 50% threshold elevation at 0.5-0.6 cpd. The similarity between these values for single neurons and for human psychophysics suggests that the acuity limit for humans is determined by the frequency at which the majority of V1 neurons fail to modulate their responses. The cutoff frequency for the neurons seems in turn to be limited by the size of the RF (Figure 4.7). Taken together these observations suggest a simple physiological explanation for the observation that cyclopean acuity for depth modulations is so poor. When the period of the corrugation is smaller than the RF size, the response of neurons with adjacent RFs becomes identical, and the corrugation cannot be detected.

## Chapter 4

# Temporal frequency tuning for disparity modulation

### 4.1 Introduction

Most physiological and psychophysical studies have focused on the spatial properties of disparity sensitivity. The temporal resolution of disparity tuned neurons –in striate or extrastriate cortex- is unknown. Given the behavioral importance of detecting motion in depth this relative lack of knowledge seems surprising. Neurons in V1 do not directly account for many aspects of stereopsis (Cumming and Parker, 1997, 1999, 2000), but see also Cumming, (2002). Nonetheless, they appear to play a critical limiting role for many aspects of behavioral performance, such as acuity for perceiving spatial modulations of disparity (Chapter 3; Nienborg et al., 2004). Here I examine the relationship between the temporal properties of disparity processing in single V1 neurons and in human observers.

A central objective of the study was to resolve a paradox raised by two existing observations. First, the temporal resolution of single V1 neurons for contrast modulation is high enough to support psychophysical performance (Hawken et al., 1996). Second, psychophysical studies have shown that the temporal resolution for disparity modulation (Norcia and Tyler, 1984) is poorer than for contrast modulation (Kelly, 1971a; Kelly et al., 1976). If the temporal frequency resolution of single V1 neurons reflects filtering characteristics intrinsic to the cell under study, then the resolution should be the same regardless of what stimulus is used to modulate the synaptic input. This suggests that

single neurons should be able to resolve disparity modulation at the same frequencies as for contrast. It follows that such neurons would signal disparity modulations over a substantial range of frequencies that are psychophysically undetectable. The cornerstone of this argument is that the temporal integration of V1 neurons should be the same for disparity changes and for contrast changes. Here, I shall examine this assumption experimentally using random dot stereograms that varied disparity as a sinusoidal function of time.

In order to compare these data with behavioral performance, I measured human psychophysical responses to disparity modulation as a function of temporal frequency, using a procedure similar to that used by Norcia and Tyler, (1984), but modified to determine disparity thresholds at each temporal frequency. This then allowed me to construct a sensitivity curve which could be compared with the behavior of V1 neurons. This combination of physiology and psychophysics thus identifies the mechanisms by which neuronal populations support multiple behavioral abilities.

## 4.2 Methods

### 4.2.1 Temporal frequency tuning for disparity modulation

I studied disparity selective neurons with random dot stimuli which varied disparity sinusoidally as described in 3.2.1. For this study of temporal responses, the spatial frequency of this modulation was low enough (usually 0.125 cpd) that the disparity at any moment was essentially uniform within a receptive field (RF). The disparity modulated between a value close to the preferred disparity and a value close to the null disparity (so that firing rate was a monotonic function of disparity within the range used). The total stimulus size was  $5^\circ \times 5^\circ$  ( $4^\circ$  modulating region +  $1^\circ$  surrounding frame of zero disparity, black and white dots,  $0.1 \times 0.1^\circ$  size, usually 99% contrast). I varied the temporal frequency of the corrugation in one octave increments (1-32 Hz for monkeys Hg, Rb, 1.125-36 Hz for monkeys Rf, Df), presented in pseudo-random order. Each stimulus condition was presented a minimum number of four times (maximum 27, mean 10.3).

For entry into this study, I required that all neurons show significant modulation of firing rate with changes in disparity (in RDS with uniform disparity) on a one-way ANOVA ( $p < 0.05$ ), and that their mean response at the preferred disparity exceeded 10 spikes/sec.

The main manifestation of disparity selectivity in response to the corrugations is a periodic modulation in firing rate at the drift frequency. I quantified this modulation by measuring the *relative modulation* (RM, defined in 3.2.3). The value for  $fl$  was obtained by averaging the responses for all trials, then calculating the modulation of this mean response at the frequency  $fl$ . The same calculation (repeated for all temporal frequencies of the corrugations) applied to responses to the planar stimulus, served as a control for the extent of modulation that was not attributable to the corrugation. In order to eliminate artifactual modulation related to response-latency and the onset transient, the first 500ms of each trial was discarded.

A resampling (1000 cycles) and bootstrapping method was used (Davison and Hinkley, 1997) to determine confidence intervals. I required that relative modulation in response to at least one temporal frequency was significantly higher (on the 5% level, corrected for multiple comparisons) than in response to the planar stimulus, before further analysis of the response modulation was undertaken.

For most cells the variation of RM as a function of temporal frequency of the corrugations gave a curve that could well be described by Gaussian functions. I therefore used Gaussian fits for subsequent analysis.

#### 4.2.2 Temporal frequency tuning for contrast

For comparison with the responses to disparity modulation I also measured temporal frequency tuning in response to drifting sinusoidal luminance gratings at high contrast (usually 99%). The stimuli were presented with the preferred orientation and spatial frequency, and over the same range of temporal frequencies as was used for disparity modulation, a minimum of four times each. Gaussian functions were fit to mean rate as a function of grating temporal frequency.

It is possible that the differences in spatial structure between the RDS and the drifting gratings might be associated with a change in temporal integration (e.g. because of changes in contrast normalization). In order to measure temporal frequency tuning for contrast in an RDS stimulus, I employed dot patterns in which the dot locations remained fixed throughout an entire 2-second trial. However, the luminance of the dots was modulated sinusoidally (as dark dots got lighter, bright dots got dimmer, until the contrast was reversed). This is equivalent to applying counterphase modulation to all of the spatial Fourier components of the dot pattern, so I describe this stimulus as a counterphase modulating RDS. The RDS were otherwise similar to the stimuli used for measuring disparity tuning. The RDS and gratings were presented monocularly, or binocularly at the preferred disparity. For complex cells the dominant response to contrast reversal is modulation at the second harmonic of the stimulus (Movshon et al., 1978b; Hawken et al., 1996). Tuning curves plotted therefore  $f^2$  (modulation amplitude at the second harmonic of stimulus temporal frequency) as a function of stimulus temporal frequency. I fitted Gaussian functions to the tuning curves.

### 4.2.3 Onset transient

Characterization of the response at onset (latency and time to 60% peak response) was obtained from averaged spike density functions. I fitted a horizontal line followed by an exponential to the onset of the spike density function (from 30ms prior to stimulus onset to the peak of the response within the first 250ms after stimulus onset):

$$f(t) = \begin{cases} b, & \text{for } t \leq l \\ ge^{-\frac{1}{\tau(t-l)}} + b, & \text{for } t > l \end{cases}$$

The fit had four free parameters: baseline,  $b$ , latency,  $l$ , gain,  $g$ , and a time constant  $\tau$ , which were constrained to values  $>0$ . Least square residuals were minimized using a non-linear optimization algorithm (Lagarias et al., 1998). The rise time to 60% peak was the

first bin in units of 0.1ms of the spike density function which was >60% of peak within the first 250ms after stimulus onset.

#### 4.2.4 Temporal integration time

To determine the temporal integration time in response to disparity modulation, the phase of the response modulation was plotted as a function of stimulus frequency. For a constant temporal integration time, the phase is proportional to the input frequency, and the slope of the resulting function corresponds to the temporal integration time (Reid et al., 1992; Hawken et al., 1996). To reduce error caused by noise, phases were only calculated at the temporal frequencies for which RM was >0.5 peak RM. The slope of the function relating phase and input temporal frequency was then determined by linear regression.

#### 4.2.5 Psychophysical experiments

I measured sensitivity to disparity modulation in four human subjects (one female, two naïve observers), with normal or corrected to normal vision. The subjects viewed the stimuli in a Wheatstone stereoscope similar to that used for recordings. Stimuli were planar dynamic random dot stereograms (size  $8^\circ \times 8^\circ$ , dot size  $0.1^\circ \times 0.1^\circ$ , 50% black, 50% white dots, 99% contrast, dot density 50%) whose disparity changed over time. In order to avoid monocularly detectable changes in the stimuli, the disparity modulating RDS was surrounded by a  $1^\circ$  frame at zero disparity. Total stimulus size was  $9^\circ$  ( $8^\circ + 1^\circ$  frame). Subjects had to discriminate a disparity modulating RDS (temporal frequency 0.5Hz to 24Hz, varying between runs) from a control RDS in a two interval forced choice procedure. The monitor framerate was set to 96Hz for the psychophysical experiments to improve the frequency resolution. The amplitude of the disparity modulation was varied pseudorandomly between trials to determine detection thresholds. Above the temporal frequency limit of disparity modulation, a disparity modulating RDS is perceived as a transparent cube (Norcia and Tyler, 1984). The control interval therefore contained an RDS giving the same percept. The disparities of the dots in the control interval were chosen such that they matched the disparities of the disparity modulating RDS (because

of the monitor frame rate the sinusoidally disparity modulating stimulus contained discrete disparity steps). For each stimulus pair, the same total number of dots was presented at each disparity in the control as in the stimulus intervals. The control interval contained the same distribution of disparities on every frame, while the stimulus interval contained only one disparity on each frame. At intermediate frequencies, this control stimulus gave rise to a thicker percept than the disparity modulating stimulus. In order to prevent subjects from using perceived thickness as a cue in the task, I scaled the range of disparities contained in the control interval to ensure that this cue could not be used.

At the lowest frequencies used ( $< 3\text{Hz}$ ), the presence of a detectable (non-zero) disparity at any point in the trial might have been used to detect the interval with modulation. At these low frequencies I therefore added a small pedestal disparity (exceeding the disparity modulation amplitude) to both stimulus and control intervals. Thus at all frequencies, subjects were forced to detect modulation in disparity over time in order to do the task. Finally, to ensure that temporal components of the onset did not distort the results, all disparities in the first 100ms were scaled down with a temporal Gaussian window, for both the control and stimulus intervals.

Subjects were asked to fixate on a cross in the center of the monitors. Each trial was initiated by a button press. The stimuli were presented foveally and usually lasted for 1 sec each (separated by a 100ms interval) unless subjects terminated trials earlier by making their choice before the end of the stimulus presentation. Cumulative Gaussian curves were fit (using a maximum likelihood estimator) to the psychometric functions to determine the point at which 75% correct performance was achieved. (I will refer to this as the disparity threshold at the respective temporal frequency of disparity modulation.)

#### 4.2.5 Gaussian fits

I used Gaussian fits to quantify temporal frequency tuning in response to disparity modulation, luminance gratings and counterphase modulating RDS, and to characterize the psychophysical performance. All fits were allowed to be a Gaussian function of linear or logarithmic temporal frequency (whichever minimized the least square residuals) in a

nonlinear optimization algorithm [based on the simplex-search method (Lagarias et al., 1998)]. The Gaussian functions had four free parameters (mean, amplitude, standard-deviation and baseline), which except for the mean were constrained to values  $\geq 0$ .

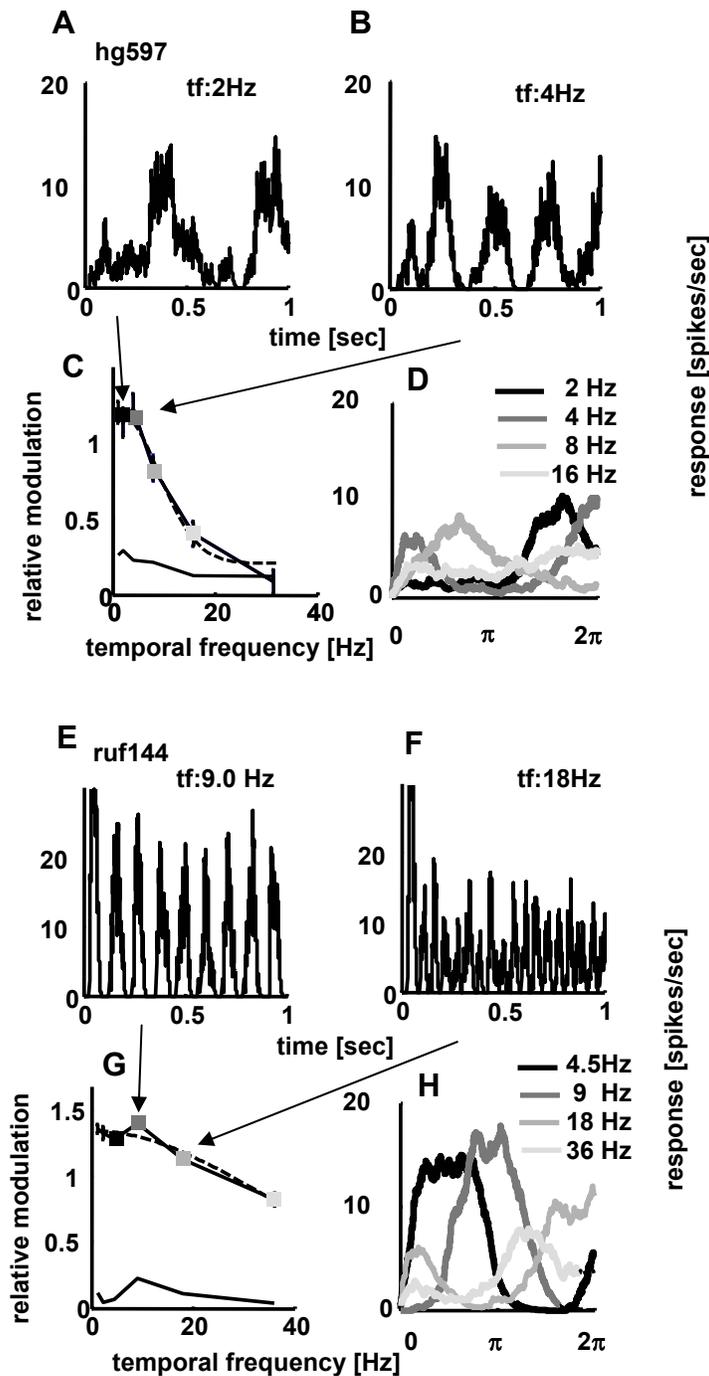
Gaussian fits were considered adequate if they explained more than 65% of the variance within the data. The temporal frequency at which the Gaussian dropped to 2/3 of its positive peak value (i.e. the peak within the segment of the Gaussian for which temporal frequencies were  $\geq 0$ ) was defined as the high cutoff frequency.

## 4.3 Results

I recorded from 242 isolated units in striate cortex of one female and three male monkeys: 110 from monkey Hg, 62 from monkey Df, 59 from monkey Rf, and 11 from monkey Rb. Of these, 117 units were significantly selective for disparity (5% level in ANOVA) and fired  $>10$  spikes/sec in response to their preferred disparity. Responses to disparity modulating RDS were analyzed for 73 of the 117 cells. (For the remainder, the unit was either lost before a complete dataset was recorded or the range of disparity-modulation did not match the range of disparity selectivity of the cell.)

### 4.3.1 Temporal frequency tuning in response to disparity modulation

During presentation of these stimuli, the disparity at any point in a receptive field modulates as a sinusoidal function of time. I have previously shown that disparity sensitive V1 neurons respond to such stimuli with a sinusoidal modulation of their response at the stimulus frequency, (Chapter 3, Nienborg et al., 2004). A neuron presented with disparity modulating RDS of different temporal frequencies should therefore change the frequency of its modulation accordingly. This is what I observed.



**Figure 4.1** Responses to disparity modulating RDS for two neurons (hg597, ruf144). Panel A,B and E,F depict spike density functions in response to disparity modulating RDS at different temporal frequencies (2Hz, 4Hz for panel A and B, and 9Hz, 18Hz for panels E and F). Panels D and H show cycle averages of the spike density functions (2Hz to 16Hz and 4.5Hz to 36Hz respectively). The relative modulation (RM, modulation amplitude at the stimulus frequency over mean firing rate) is plotted as a function of

temporal frequency in panel C (cell hg597) and G (cell ruf144). Baseline RM is the modulation at each temporal frequency in response to a planar control stimulus (no modulation). The arrows point to the RM calculated for the data depicted in the respective panels A,B and E,F. RM plotted as a function of temporal frequency is fitted by Gaussian curves (dashed lines in panels C and G). The temporal frequency high cutoff was calculated at  $2/3$  (peak RM – mean baseline RM). High cutoff values were usually similar to that shown in panel C (8.1Hz) but were significantly higher in a few neurons, as seen in panel G (35.3Hz).

Figure 4.1 (panels A, B and E, F) depicts the spike density functions of two cells in response to disparity modulating RDS of different temporal frequencies. Panels D and H show cycle averages of these spike density functions allowing clearer comparison of the response amplitude and phase. These neurons, like all of those I observed, modulated their firing at the stimulus frequency. I quantified the extent of modulation by calculating the relative modulation (RM, Movshon et al., 1978a). It is the amplitude of the neuronal modulation at the temporal frequency ( $f_1$ ) of the stimulus divided by the mean firing rate of the neuron ( $f_0$ ). In order to control for modulation unrelated to the stimulus modulation I also measured the response to a corrugation whose depth modulation was 0, i.e. a planar stimulus with a constant disparity over space and time, interleaved with the modulating stimuli. In 69 out of 73 neurons, the modulation was significantly stronger in at least one of the temporal frequencies tested, than in response to the planar control stimulus (by resampling).

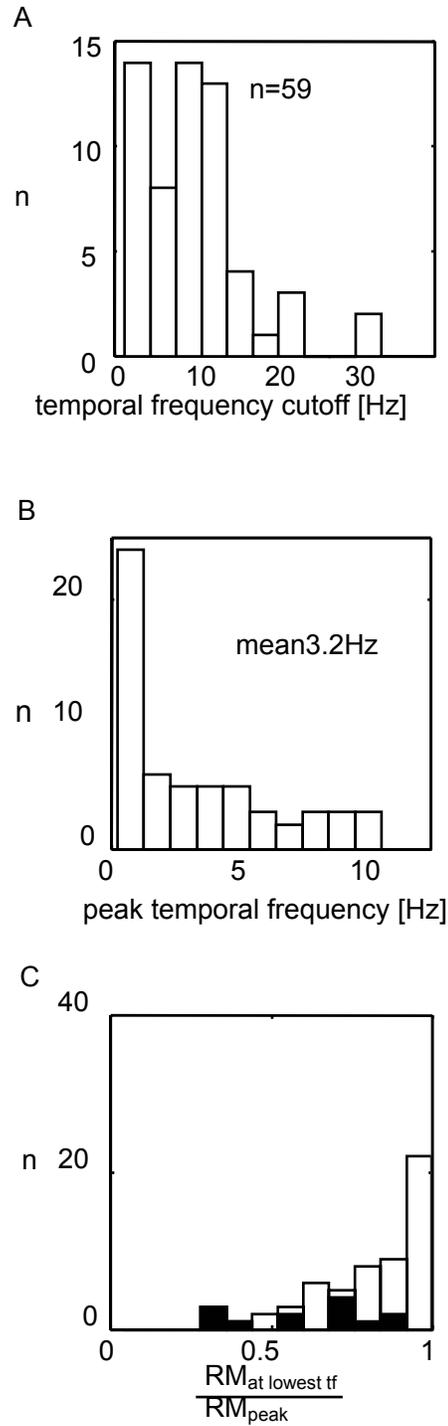
The mean firing rate was little affected by temporal frequency. Only 32/69 cells showed any significant change (one-way ANOVA,  $p < 0.05$ ), and even for these cells the modulation in mean firing was modest and rarely showed systematic changes as a function of temporal frequency. I therefore used the RM-values to obtain temporal frequency tuning curves in response to disparity modulation (Figure 4.1C and 4.1G), and described these by fitting Gaussian curves. For 59/69 cells, the Gaussian fits explained  $>65\%$  of the variance and were used to describe the temporal properties [high-cut temporal frequency (frequency at  $2/3$  peak RM), peak frequency and frequency bandwidth obtained from the mean and the S.D. of the Gaussian fit respectively].

Unless stated otherwise the remainder of the quantitative analysis is restricted to these 59

cells. The mean of the peak frequencies was  $3.2\text{Hz} \pm 2.9\text{Hz}$  (Figure 4.2B). The mean high temporal frequency cut off was  $10.0\text{Hz} \pm 6.5\text{Hz}$  S.D. (Figure 4.2A), and the mean of the frequency bandwidth was  $6.1 \pm 6.7\text{Hz}$  (data not shown). The distribution of the peak temporal frequencies shows that a large number of cells have peaks at 1Hz (the lowest frequency tested), suggesting that they are lowpass. In order to quantify the extent to which cells were lowpass/bandpass I calculated the ratio of RM at the lowest temporal frequency over peak RM (I obtained the values for this ratio from the raw RM data, not from the fit). Figure 4.2C depicts the distribution of this ratio, which is strongly biased towards values of 1. Nonetheless, 13/59 cells (filled bars) showed statistically significant low frequency attenuation. ( $p < 0.05$ , two tailed test by resampling, corrected for multiple comparisons). The magnitude of the attenuation was generally modest, with only 4/59 neurons showing responses that fell below half their peak value.

It is possible that even these figures overestimate the true magnitude of low frequency attenuation, if the disparity modulation induced vergence eye movements. Monkeys trained to make tracking vergence eye movements are able to do so only up to frequencies of 1-4Hz (Cumming and Judge, 1986), depending on stimulus amplitude. If such movements were elicited by the stimuli used here, they would have reduced modulation in retinal disparity over the RF of the neuron. Since this reduction would only occur at low stimulus frequencies, it could lead to the appearance of low frequency attenuation. However, examination of vergence records revealed no systematic responses to the disparity modulation in the stimulus. The measured amplitude of vergence modulation was on average only 15% of the modulation in stimulus disparity. Three observations suggest this vergence variation largely reflects artifacts [as also suggested by Read and Cumming, (2003a)]. First, the amplitude of the response at any one frequency was independent of the stimulus modulation frequency. Second, the extent of modulation observed in vertical vergence was similar to that seen for horizontal vergence, despite the fact that there was no modulation in the vertical disparity of the stimulus. Finally, I found no correlation between the relative amplitude of the vergence eye movements and the extent of low-frequency attenuation. Thus it appears that the

reduction in neuronal response amplitude at low stimulus frequencies is a real property of the neuron's response to disparity, not an artifact related to vergence eye movements.



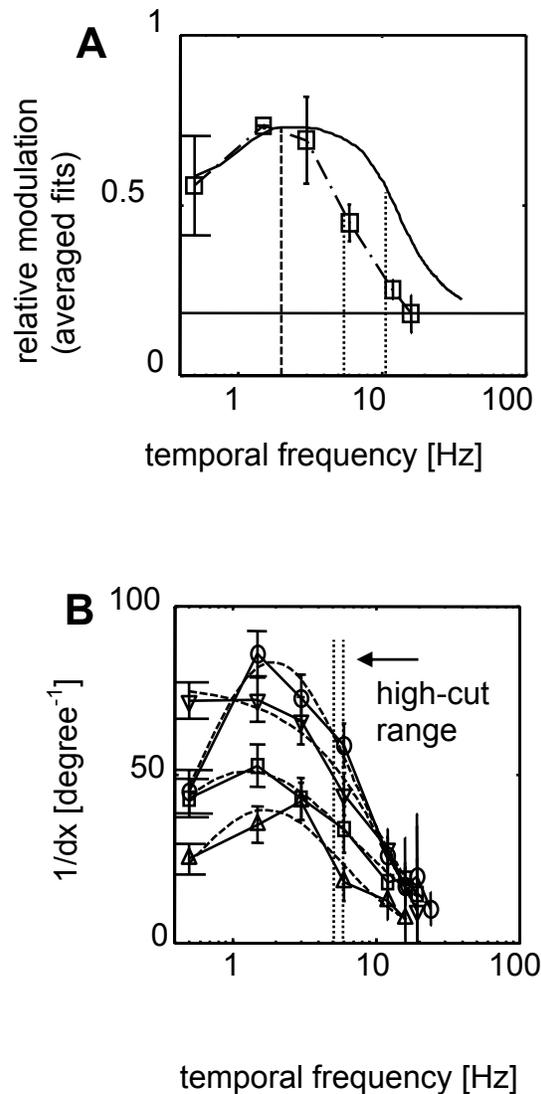
**Figure 4.2** Temporal frequency tuning of the cell population. Panel A summarizes the high frequency cutoff for 59 neurons. Mean value is 10.0Hz. A frequency histogram of

the peak temporal frequencies of the 59 neurons is shown in panel B (mean 3.2Hz). Most neurons had low peak temporal frequencies. The ratio of RM at the lowest temporal frequency over peak RM was used to estimate the extent of low frequency attenuation (frequency histogram in C, n=59). Filled bars correspond to cells with statistically significant low frequency attenuation (n=13,  $p < 0.05$  by resampling).

To summarize the responses of the whole neuronal population, I averaged the Gaussian fits (n=62) describing the variation of RM as a function of temporal frequency (Figure 4.3A). [Note that this average includes the fits for cells whose modulation was not significantly stronger in response to disparity modulation than in response to the planar control stimulus (n=3).] The temporal frequency cutoff for this averaged fit (solid line in Figure 4.3A) was 10.5Hz (dotted line in Figure 4.3A), similar to the mean of the temporal frequency cutoffs, and the peak was 2Hz (dashed line in Figure 4.3A). The low frequency roll-off was modest. If the temporal integration of disparity signals in striate cortex is a limiting factor for psychophysical performance, these features might be reflected in the psychophysical sensitivity to disparity modulation. A previous study showed that the highest temporal frequency at which humans were able to detect motion in depth in RDS was 9Hz, and to detect flickering in depth 14Hz (Norcia and Tyler, 1984). However, it is difficult to relate these measures of limiting frequency with the continuous sensitivity curve I show for the neuronal population (Figure 4.3A). I therefore used a protocol very similar to that of Norcia and Tyler, but measured disparity thresholds at each frequency. The reciprocal of these thresholds then yields a sensitivity curve for human observers that can be compared with the neuronal population (Figure 4.3B).

Comparing the neuronal and psychophysical sensitivity curves, there are several close similarities. First, despite differences between subjects in absolute sensitivity, the temporal frequency high cut off, (defined as 2/3 peak performance) was almost identical in all subjects (5.1Hz to 5.9Hz, mean 5.5Hz $\pm$ 0.4Hz S.D.). This is similar to the equivalent measure for the neuronal population, although slightly lower. Second, for all four subjects performance was best between 1.5Hz and 3Hz which corresponds to the peak of the neuronal population (2Hz). Third, the extent of low-frequency attenuation is similar for the psychophysical and neuronal data. These similarities between psychophysical and

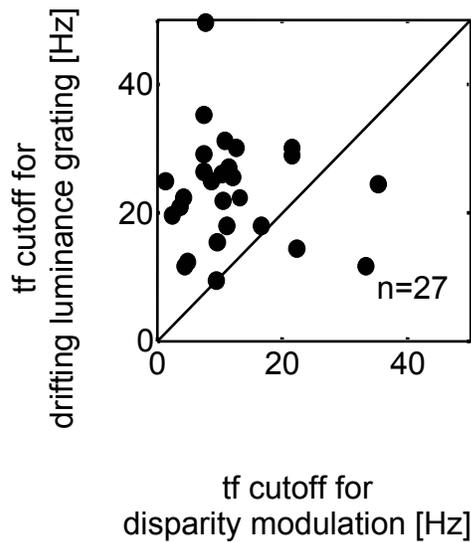
neuronal data strongly suggest that the initial calculation of disparity in striate cortex is what limits the psychophysical temporal resolution for stereopsis.



**Figure 4.3** Comparing the neuronal population response with the psychophysical performance. Panel A depicts the averaged Gaussian fits to the RM as a function of temporal frequency for the 62 neurons (solid line). The baseline is the mean RM in response to the planar control (no modulation) for these 62 neurons. Superimposed is the mean of the psychophysical performance (normalized for each subject by the value at 1.5Hz) in response to disparity modulation for four human subjects (open squares). The dotted vertical lines indicate the high cutoff for the averaged fits of the neuronal data (10.5Hz) and the mean human psychophysical performance (5.5Hz). Psychophysical performance for each subject (n=4) is shown in panel B. Sensitivity (1/disparity threshold, see Methods) is plotted as a function of temporal frequency. The range of the temporal frequency cutoffs between subjects is indicated by the dotted lines.

What makes this conclusion surprising is that psychophysical detection of temporal modulation signaled by contrast is very different from the equivalent measures for disparity- despite both apparently being limited by the temporal resolution of V1 neurons. Sensitivity to contrast modulation is a bandpass function of frequency (Kelly, 1971b), as are the responses of many V1 neurons (Hawken et al., 1996). More strikingly, an observer's highest detectable temporal frequency for contrast modulation is substantially higher than for disparity. Under optimal conditions (high luminance, large fields) human flicker sensitivity can reach 80Hz (Kelly, 1961). For smaller stimulus sizes similar to those used in recording experiments (2-4degrees) and otherwise optimal conditions, contrast sensitivity falls to 2/3 peak at about 30Hz (de Lange, 1958; Keeseey, 1970; Kelly, 1971b), which is significantly higher than the temporal frequency cut off I found in response to disparity modulation. Yet this better high frequency performance is matched by the properties of V1 neurons: the mean high-cut and peak frequency to luminance gratings reported for V1 neurons (Hawken et al., 1996) are comparable to these psychophysical data, and are approximately 2.5 times the values I measured for neuronal responses to disparity modulation. This comparison raises the question of how V1 neurons are able to respond at higher temporal frequencies in response to contrast modulation than in response to disparity modulation.

One possibility is that different populations of neurons were studied. In V1 only about 60% of the neurons are disparity selective (Prince et al., 2002b). If disparity tuned neurons tended to prefer lower temporal frequencies than the average V1 population, this discrepancy would simply be the consequence of my sampling only disparity tuned neurons. To address this possibility, I directly compared the high-frequency cut off in response to disparity modulation with that in response to drifting luminance gratings, in the neuronal population.



**Figure 4.4** Temporal frequency cutoffs in response to drifting luminance gratings is generally higher than in response to disparity modulation. Temporal frequency cutoffs in response to disparity modulating RDS were compared with that in response to drifting luminance gratings for 27 neurons. Only three neurons lie statistically significantly below the identity line (by resampling). The ratio of the cutoff in response to drifting luminance grating over cutoff in response to disparity modulation is significantly larger than 1 ( $p < 0.001$ , geometric mean 2.3).

In Figure 4.4, the temporal frequency high cut in response to drifting luminance gratings is plotted against the high cut in response to disparity modulation for 27 cells. Only three cells lie significantly below the identity-line (by resampling). The mean temporal frequency high-cut for the gratings is 23.4Hz, similar to the values reported by Hawken et al. (1996). The correlation is not significant ( $r = -0.09$ , n.s.). The geometric mean ratio of grating cut-off over disparity cutoff is 2.3, which is significantly larger than one ( $p < 0.001$ , by resampling). Thus it appears that the temporal frequency tuning of single neurons is different for the two stimuli.

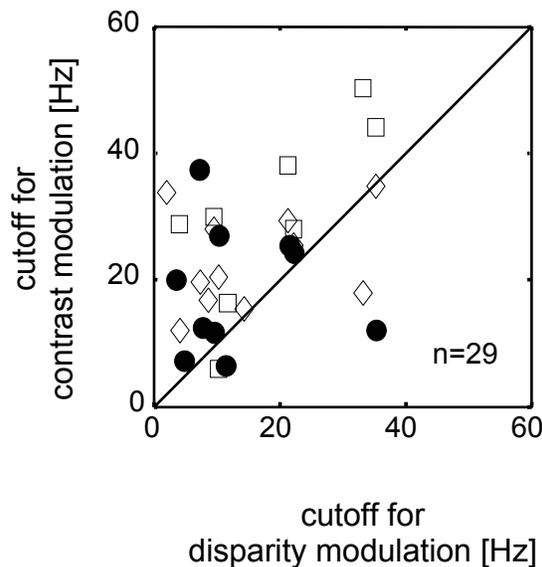
This is different from the changes in temporal frequency tuning known to occur with changes in contrast (Holub and Morton-Gibson, 1981; Albrecht, 1995; Carandini et al., 1997), where the changes probably reflect the power of the stimulus for driving the cells.

I used full contrast both for the gratings and the RDS. Mean firing rates did not vary systematically between the two stimuli for the cells in Figure 4.4 (the geometric mean ratio of rates in response to the RDS over rates in response to the gratings was 0.9, n.s. different from 1, by resampling). This suggests that both stimuli were similarly effective in driving the cells and that different driving power of the stimuli cannot explain the difference in temporal frequency tuning.

When trying to understand the origin of these different temporal properties remember that the temporal frequency tuning curves for disparity modulation were based on RM. In contrast, the temporal frequency cut offs in response to drifting luminance gratings were obtained from mean firing rates. Suppose a cortical cell receives input from LGN-neurons that can modulate their response up to higher frequencies than the cortical cell [as results by Hawken et al. (1996) suggest]. Then the cutoff frequency for modulation in firing rate would be limited by the temporal properties the cortical neuron. But at higher temporal frequencies the cortical neuron would still receive modulating input from the LGN, which may be sufficient to maintain a mean firing rate above baseline.

Following this reasoning, one possible explanation of the results might be that some input elements of the cells in Figure 4.4 have higher temporal frequency cut-offs than the output (as measured by the RM). Such an explanation should hold independently of what stimulus is used to drive response modulation. I therefore examined changes in firing rate elicited by contrast modulating stimuli. For simple cells ( $F1:F0 > 1$ ), I examined the F1 component in response to drifting gratings. In complex cells the response to drifting gratings is dominated by an increase of discharge relatively independent of the spatial phase of the stimulus (Movshon et al., 1978b). In order to modulate the response rates of complex cells, I used counterphase modulating stimuli (gratings or RDS). The dominant response of complex cells is modulation at the second harmonic of the stimulus frequency (Movshon et al., 1978b; Hawken et al., 1996). The use of counterphase modulating RDS also ensures that any changes in temporal integration cannot be explained by differences in contrast normalization secondary to changing the spatial properties of the stimulus. Figure 4.5 compares these measures of high frequency cutoff

for each neurons' response modulation with that observed in response to disparity modulation. Note that the abscissa marks the frequency of the cell's modulation. For the open symbols (counterphase modulating stimuli) it therefore corresponds to double the stimulus frequency. As in Figure 4.4, the cutoff frequency is systematically higher for contrast modulation, with only two cells significantly below the identity line. The geometric mean ratio of the temporal frequency cutoff in response to contrast modulation over that in response of disparity modulation is 1.8, significantly larger than 1 ( $p < 0.001$ , by resampling).



**Figure 4.5** Temporal frequency cutoff for response modulation driven by contrast is significantly higher than in response to disparity modulation. Ten cells showed clear modulation at the stimulus frequency in response to drifting gratings ( $F1 > F0$ , simple cells). For these cells the cutoff frequency for the  $F1$  component is plotted (filled circles). For the remaining cells, the responses to counterphase modulating stimuli RDS (open diamonds,  $n = 11$ ) or gratings (open squares,  $n=8$ ) were analyzed. Note that the abscissa plots the frequency of the modulation in neuronal firing (i.e. twice the stimulus frequency for the open symbols). They are plotted against temporal frequency cutoffs in response to disparity modulation. The geometric mean ratio (cutoff in response to contrast modulation over cutoff based on disparity modulation) is significantly higher than one (1.8,  $p < 0.009$ , by resampling).

This indicates that the neurons are able to modulate their response rates at higher frequencies in response to contrast modulation than in response to disparity modulation. This cannot be explained simply by suggesting that cortical neurons modulate their output more sluggishly than their inputs. In addition, comparing the modulation amplitudes in response to disparity and contrast modulation yielded similar values on average (the geometric mean ratio of the maximal amplitude for disparity and contrast modulation was 1.0, n.s. from 1). This suggests that the lower temporal frequency cut off does also not result from a decrease in signal to noise ratio for disparity modulation.

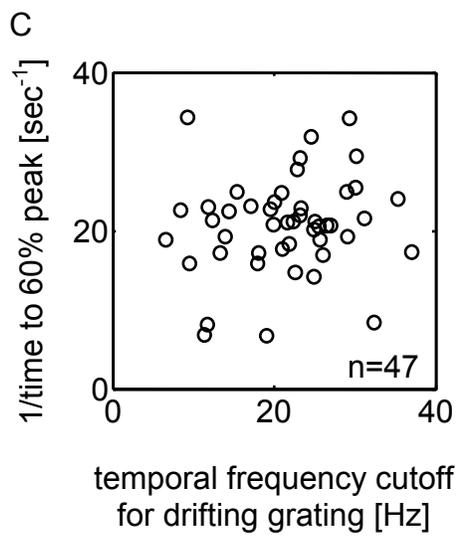
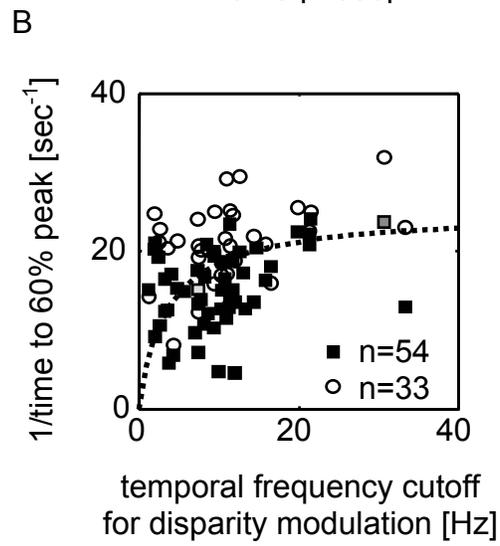
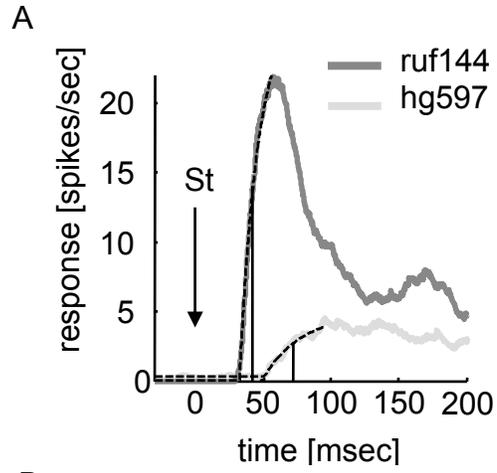
Rather, the lower temporal frequency cutoff seems to be specific to the disparity input driving the modulation. Studies of temporal responses in the LGN (Purpura et al., 1990; Hawken et al., 1996) and monocular temporal-spatial receptive field maps (Cai et al., 1997; Anzai et al., 2001; Menz and Freeman, 2004) suggest that the monocular images are temporally bandpass-filtered. For contrast modulation, the response therefore reflects the frequency response of the band-pass filter. Simple models that successfully describe many features of disparity selective neurons in V1 (e.g. binocular energy model) functionally compute cross-correlations of monocular images (Fleet et al., 1996; Qian and Zhu, 1997; Anzai et al., 1999a).

In order to understand the behavior of such a model to my stimulus, I explored the effect of temporal filtering of the monocular images on the calculation of the correlation. Appendix C shows that when the monocular images are temporally bandpass filtered, subsequent measures of binocular correlation have a frequency response which is very different from that of the filter. The response to changes in correlation has a low-pass response (as observed for the majority of the neurons, Figure 4.2C). For similar monocular temporal kernels, the frequency response of the correlation is determined by the Fourier transform of the square of the temporal kernels. Because of the squaring, the temporal frequency cutoff is largely determined by the width of the monocular temporal filters – regardless of their temporal structure within it. The lower temporal frequency cutoff for disparity modulation than for contrast modulation therefore follows directly from the computation of correlation of the monocular inputs by the binocular neuron.

### 4.3.2 Further examination of the temporal properties of the neurons: response onset and -phase

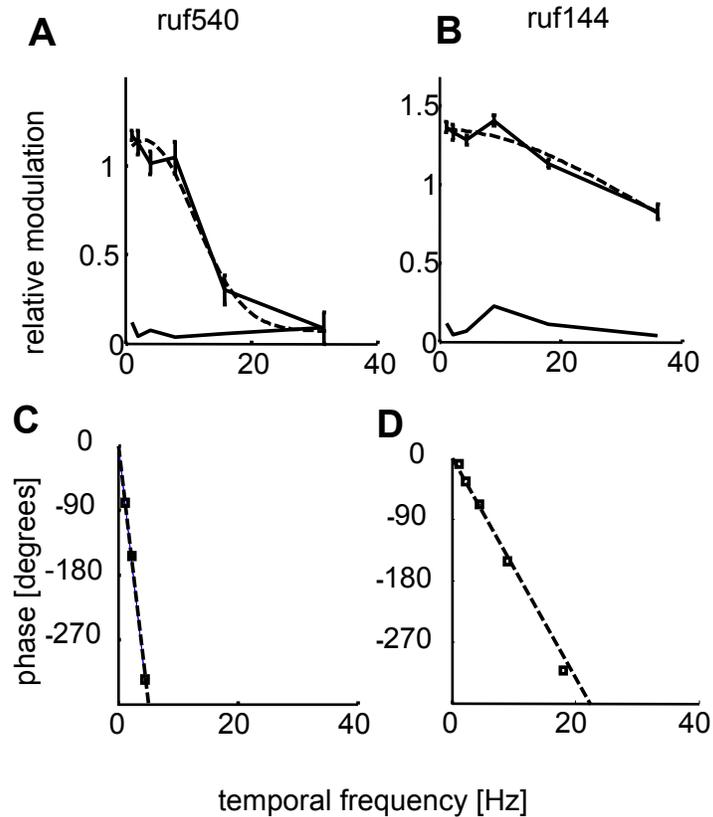
The above scheme also predicts a reciprocal relationship of the temporal frequency cutoff for disparity modulation with two other properties: 1) the onset rise time (i.e. the first point in time after stimulus onset when the response reached 60% its peak, Figure 4.6A and see Methods); and 2) the integration time deduced from phase lags.

Figure 4.6B shows that neurons with faster initial onsets (shorter times to 60% peak response) were able to modulate at higher temporal frequency ( $r=0.35$ ,  $p<0.01$ ,  $n=54$ , filled squares). The relationship can be described by a prediction (dashed line in Figure 4.6B) obtained for a fixed delay of 40ms, followed by a computation of correlation between the temporally bandpass filtered monocular images (the temporal band-pass kernel was chosen to be consistent with physiological data reported for the LGN, Purpura et al., 1991; see also Appendix C). Despite an overall shift towards shorter rise times the correlation was also significant when the rise time was obtained in response to drifting luminance gratings (Figure 4.6B, open circles,  $n=33$ ,  $r=0.42$ ,  $p<0.02$ ). I also found a significant correlation between the temporal frequency cutoff in response to disparity modulation and the reciprocal of the latency ( $r=0.4$ ,  $p<0.01$ ,  $n=41$ , data not shown). The correlation reflects the fact that response latencies in visual cortical neurons are in the order of ten times larger than expected from pure conduction delays (Reid et al., 1992), due to temporal filtering in addition to conduction delays. The temporal frequency high cutoffs in response to drifting luminance gratings was not correlated with the rise times in response to these stimuli (Figure 4.6C,  $r=0.14$ ,  $p>0.3$ ,  $n=47$ ). This agrees with results by Hawken et al. (1996), who reported no correlation between their measures of temporal integration time and temporal frequency high cutoff in response to luminance gratings in V1.



**Figure 4.6** Temporal frequency cutoff to disparity modulation and rise time at response onset. Panel A shows averaged spike density functions of the onset in response to disparity modulating RDS. The arrow points to the onset of the stimulus (St). Dashed lines depict fits used to determine response latency and rise time to 60% peak (see Methods). Solid lines mark rise time to 60% peak. The neuron (ruf144) with the shorter time to 60% peak (42ms) has the higher temporal frequency cutoff (35.3Hz). Neuron hg597 has a time to 60% peak of 72ms and a temporal frequency cutoff of 8.1Hz. (The cells are the same for which the spike density functions in Figure 4.1 are shown.) The scatterplot in panel B compares the temporal frequency cutoffs in response to disparity modulation with the reciprocal of the time to 60% peak obtained for disparity modulating RDS (n=54, filled squares) and drifting luminance gratings (n=33, open circles). The correlation is significant ( $r=0.45$ ,  $p<0.002$  and  $r=0.47$ ,  $p<0.01$  respectively) as expected for a cross-correlation of bandpass filtered images (dashed line, see Results and Appendix C). No correlation is found between the time to 60% peak for drifting gratings with the highcuts in response to these stimuli (Panel C, n=47,  $r=0.14$ , n.s.).

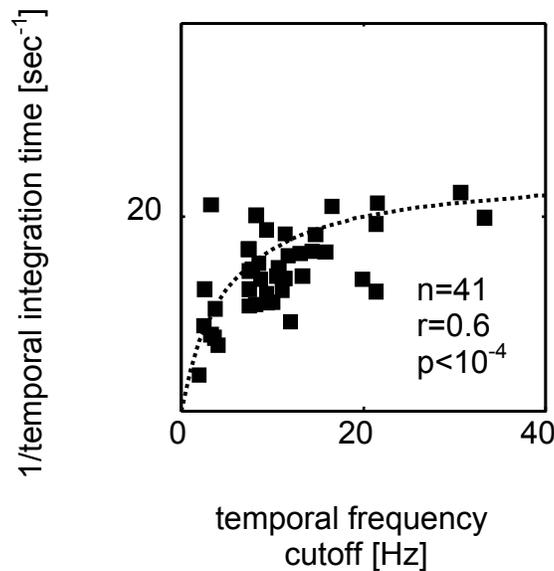
Appendix C shows that measuring the correlation of temporally bandpass filtered images is equivalent to filtering with the squared monocular filter (which is similar to the square of the monocular envelope, i.e. to a lowpass filter). For a low-pass filter, the phase lag increases systematically with frequency. The slope of this relationship can be used to estimate the temporal integration time (Reid et al., 1992, Hawken et al., 1996). It corresponds to the sum of the conduction delay and the delays caused by the temporal filtering.



**Figure 4.7** Phase of the neuronal responses. Panels A and B depict two temporal frequency tuning functions in response to disparity modulation, depicting RM (ordinate) as a function of temporal frequency (abscissa). The temporal frequency cutoffs are 10.4Hz (cell ruf540, panel A) and 35.3Hz (cell ruf144, panel B). In panels C and D the phase of the averaged peri-stimulus time histogram (PSTH) is plotted as a function of temporal frequency for the cells in A and B respectively. The slope of the line relating phase and temporal frequency (dashed line) was obtained by linear regression. It will be referred to as temporal integration time (206ms and 44ms for ruf540 and ruf144, respectively). Note the longer temporal integration time (steeper slope) for the neuron with the lower temporal frequency cutoff (ruf540 panel, A and C), and the intercepts of the dashed line with the ordinate at about 0 for both cells.

Figure 4.7 (bottom) plots phase as a function of temporal frequency for two cells. (For comparison, the variation of RM as a function of temporal frequency for the cells is shown in the upper row.) In order to reduce the error caused by noise I calculated phases only for RM-values  $> 0.5$  peak RM. Note that the slope of the line relating phase and temporal frequency (dashed line in Figure 4.7C and 4.7D) is steeper (i.e. that the cell has a longer temporal integration time) for the cell (ruf540) with the lower temporal

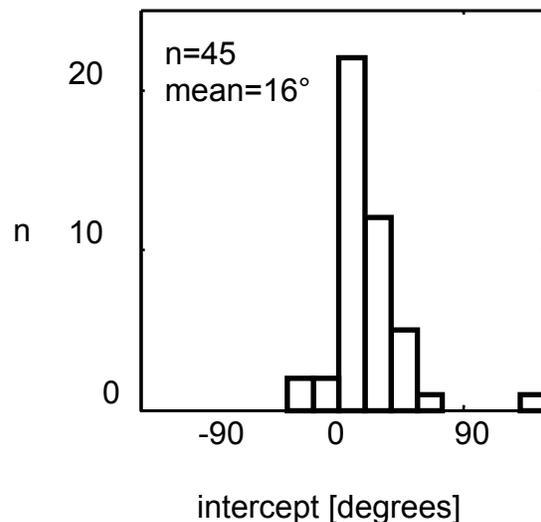
frequency cut off. I computed temporal integration times for all cells for which I had at least three phase-values. The mean temporal integration was  $72\text{ms} \pm 23\text{ms}$  S.D., very similar to the values reported for complex cells measured in response to contrast modulation ( $73\text{ms}$ , Hawken et al., 1996). The correlation of the reciprocal of the temporal integration time with temporal frequency cut off in response to disparity modulation is highly significant (Figure 4.8,  $r=0.6$ ,  $p<10^{-4}$ ,  $n=41$ ). As in Figure 4.6B, a prediction obtained for a constant delay of  $40\text{ms}$ , followed by a squared bandpass filter (which is similar to a lowpass filter, see Appendix C) provides a good description of the data (dashed line). Taking the data from Figures 4.6 and 4.8 together, it appears that all aspects of the temporal response to disparity changes in RDS are compatible with temporal bandpass filtering followed by measuring the correlation between the filtered images. This results in changes of correlation which are equivalent to low-pass filtering.



**Figure 4.8** Temporal integration time and temporal frequency cutoff. Temporal frequency cutoffs were significantly correlated with the reciprocal temporal integration time (the slope of the line relating response phase and temporal frequency, see Figure 4.7C, D),  $n=41$ ,  $r=0.6$ ,  $p<10^{-4}$ . The mean temporal integration time is  $72\text{ms} \pm 23\text{ms}$  (S.D.). The dashed line depicts the prediction for a constant delay of  $40\text{ms}$  followed by cross-correlation of monocularly band-pass filtered input.

### 4.3.3 No specialization for signaling motion in depth

Plotting phase as a function of temporal frequency also allowed me to examine whether the neurons were specialized for signaling motion in depth. A previous study found a small number of macaque V1 neurons sensitive to opposite directions of image motion in the two eyes (Poggio and Talbot, 1981), and suggested that some V1 neurons are tuned for motion in depth. However, as Maunsell and Van Essen, (1983) pointed out, when using targets that move in depth, care must be taken that changes in mean disparity alone are not responsible for the tuning observed. These difficulties are all avoided in my analysis where only the temporal frequency of modulation varies. Sensitivity for motion in depth should result in an additional constant phase shift independent of stimulus frequency. The intercept of the line relating phase and stimulus frequency would then be at values close to  $\pm 90^\circ$ . In the examples in Figure 4.7 this is not the case. Figure 4.9 shows the distribution of this intercept for 45 cells. The mean is  $16 \pm 28^\circ$ . This is close to 0 and only 1 cell has an intercept outside the range of  $\pm 65^\circ$ . Thus V1 neurons appear insensitive to the rate of change of disparity within the receptive field.



**Figure 4.9** The intercept of the line relating phase of the response and temporal frequency with the ordinate. The intercept (in degrees) with the ordinate in the phase plots (Figure 4.7C, D) is shown in the frequency histogram for 45 cells. Most values are close to 0 degrees (mean  $16.0 \pm 28.2^\circ$ SD) suggesting that the neurons respond to the instantaneous disparity and not to motion in depth.

# Chapter 5

## Discussion

### 5.1 Summary of the main results

- 1) I found that most disparity selective neurons in striate cortex responded to a sinusoidal corrugation with a lowpass function of spatial frequency (corrugation frequency). This suggests that the neurons respond strongest to a uniform disparity within their receptive field. My analysis shows that the neurons have fronto-parallel planar receptive fields, no specialization for signaling slant and no center-surround interaction. In such a scheme the high corrugation frequency cutoff is limited by the receptive field size. My independent measurements of receptive field size were compatible with this.
- 2) All of this behavior closely matched the model simulations based on the binocular energy model (functionally corresponding to binocular cross-correlation) for disparity modulations.
- 3) The high cutoff corrugation frequency of V1 neurons is similar to the acuity limit for detecting disparity modulation of human observers. V1 receptive field size may therefore impose an initial acuity-limit on processing depth, just as retinal photoreceptors do on processing shape.
- 4) I found that the temporal frequency high-cut in response to disparity modulation for V1 neurons is on average 10.5Hz. This value is similar to the temporal frequency high cut for detecting disparity modulation that I measured in human

- observers. This suggests that disparity selective V1 neurons also limit the temporal resolution for detecting disparity modulation.
- 5) The temporal frequency highcut of the neurons in response to disparity modulation was on average lower than that in response to drifting luminance gratings. Additional measurements suggest that this cannot simply be explained by differences in the analysis (temporal frequency measurements based on response modulation or based on mean responses). Rather the different temporal tuning follows mathematically as a result of calculating cross-correlation between temporally bandpass filtered monocular images. It is therefore a necessary feature of simple models for disparity selectivity in V1 which functionally correspond to a cross-correlation of monocular images, such as the binocular energy model.
  - 6) My analysis showed that disparity selective V1 neurons are not specialized for signaling motion in depth.

## 5.2 Three-dimensional receptive field structure of V1 neurons and spatial acuity for detecting disparity modulation

Previous studies of disparity selectivity in the striate cortex have used stimuli which correspond to planar fronto-parallel surfaces. This differs markedly from studies of responses to contrast, where the two-dimensional structure of the stimulus (orientation, spatial frequency etc.) is more important than its luminance. In order to explore the three-dimensional structure of V1 receptive fields, I examined responses to sinusoidal modulation of disparity in random dot stereograms. Every aspect of the response I examined could be explained as a response to the weighted mean of the disparities of the dots covering the RF (the weights being set by the RF envelope), as if the optimum disparity is the same at all locations within the RF

First, this predicts that the response should be a sinusoidal modulation in activity, at the temporal frequency of the stimulus, exactly as I observed. Second, the extent of modulation (*relative modulation*, RM) should be greatest at low frequencies. This was the typical pattern. Third, the way in which the RM declined with increasing stimulus

corrugation frequency was well described by a Gaussian function of frequency. This is the form expected if the shape of the RF envelope is also a Gaussian. Furthermore, this cutoff frequency never significantly exceeded the limit predicted from an independent measure of the RF envelope. Finally, the correlation between disparity selectivity (DDI) and the maximum value of RM indicates that there is not a subpopulation of neurons that are weakly tuned to disparity but strongly tuned for disparity gradients.

All of these properties of V1 neurons were readily matched with an implementation of the energy model (Ohzawa et al., 1990), extended to include two dimensional RFs in each eye (Bridge et al., 2001). This therefore represents another success for the most widely accepted mechanism by which disparity selectivity arises in V1. It is important to note that, taken by themselves, the data reported here are compatible with a wide range of other models. Any model in which the response is on average determined by the mean disparity of the dots weighted according to their location in the RF would suffice.

The general absence of attenuation in RM at low frequencies excludes two plausible possibilities for an underlying RF structure. First it indicates a lack of any center-surround organization in the disparity domain (i.e. any suppression from the surround does not depend on disparity, at least over the range of disparities tested here). If a region activated by one disparity was surrounded by a region where the same disparity suppressed the response, then a uniform field of dots at that disparity would produce a weaker response than a stimulus with an intermediate corrugation frequency placing the preferred disparity at the center but not in the surround. A few neurons (7/55) showed weak but significant attenuation of their response at low corrugation frequencies. The phase of their responses was compatible with a weak center-surround organization of this type. However, the attenuation was modest even in this small subpopulation, so it does not appear to be a significant feature of the response of the population of V1 neurons.

Second, the lack of low frequency attenuation argues against a specialization for signaling slant. These depth corrugations place a time-varying slant over the receptive field, and since the amplitude of the corrugations was constant, the magnitude of the slant

variation increased with the corrugation frequency. A neuron activated only by non-zero values of slant would cease responding at low frequencies. Note that this observation does not exclude the possibility that the activity of some neurons is weakly affected by stimulus slant. It does indicate that the effects of changes in slant are small compared with the effects of changes in mean disparity. This is a feature of the energy model, even when inter-ocular differences in preferred orientation are incorporated (Bridge et al., 2001).

One possibility is that because I only used disparity modulations in one orientation, I might have missed a population of neurons that are selective for tilt about other axes. If this were the case, the population would have to be strongly biased away from responding to tilt around a horizontal axis to explain the lack of slant selective responses I observed, which seems unlikely. To measure responses to different corrugation frequencies at multiple orientations would require substantially more data per neuron. Furthermore, disparity modulations presented at any other orientation introduce monocular changes to the stimulus (variation in dot density) that depend upon spatial frequency and orientation. Therefore, had I found neurons that appeared slant selective only for variation in disparity along non-vertical axes, this would not have provided clear evidence for slant selectivity.

The lack of slant selectivity in V1 raises the question of how slant selectivity in extrastriate cortical areas arises. The most natural possibility is that this is derived by combining the outputs of V1 disparity selective neurons. This allows a neuron to detect slant by means of piecewise planar (fronto-parallel) approximations. Studies of slant selectivity in extrastriate cortex have all used large enough stimuli that this explanation is entirely possible (Shikata et al., 1996; Sakata et al., 1999; Tsutsui et al., 1999; Sugihara et al., 2002; Nguyenkim and DeAngelis, 2003). Furthermore a recent study of slant selectivity in area MT (Nguyenkim and DeAngelis, 2003) found responses to pairs of small planar patches that were commensurate with this explanation. One consequence of such an explanation is that modulations over regions as small as V1 RFs should not be detected by such neurons. That is to say that if tested with depth corrugations like those

used here, this model predicts that high-frequency cutoffs for slant selective neurons should not be higher than those reported here.

Although such responses have not been studied in extrastriate cortical areas, one can pursue the same logic all the way to psychophysical measures. If slant selectivity throughout the brain is derived from piecewise comparison of outputs from V1 neurons, then the high-frequency cutoffs reported here should limit the psychophysical observer's ability to detect disparity modulations at high corrugation frequency. Human psychophysical studies (Tyler, 1974; Schumer and Ganz, 1979; Howard and Rogers, 1995; Bradshaw and Rogers, 1999; Banks et al., 2004a) have found that the acuity to detect disparity modulations is surprisingly poor (when considering the high acuity to detect luminance structure or stereo-acuity). Banks and colleagues, (Banks et al., 2004a) show that this cannot be explained by limitations inherent to the stimulus and also argue that using piecewise frontal disparity estimates imposes a limit on the highest attainable acuity for disparity modulations. Comparing the physiological data here with psychophysical responses at similar eccentricities suggests that spatial integration of these V1 neurons is what limits performance. The neuronal data show a reduction to 66% of maximal sensitivity at a mean frequency of 0.5 cpd, at a mean eccentricity of 3.7°. If psychophysical judgments depended on these responses, a corrugation frequency of 0.5 cpd should produce thresholds 50% higher than the optimum. The two human studies that examined thresholds at these eccentricities (Prince and Rogers, 1998; Banks et al., 2004a) found 50% threshold elevation at 0.5-0.6 cpd. The similarity between these values for single neurons and for human psychophysics suggests that the acuity limit for humans is determined by the frequency at which the majority of V1 neurons fail to modulate their responses. The cutoff frequency for the neurons seems in turn to be limited by the size of the RF (Figure 4.7). Taken together these observations suggest a simple physiological explanation for the observation that cyclopean acuity for depth modulations is so poor. When the period of the corrugation is smaller than the RF size, the response of neurons with adjacent RFs becomes identical, and the corrugation cannot be detected, as also suggested by recent psychophysical measurements of correlation-thresholds (Banks et al., 2004b).

These comparisons of human psychophysics with neuronal responses in the monkey of course rely on the assumption that psychophysical measures of the same phenomena in monkeys would give rise to similar results. As most studies of stereoscopic visual function in the monkey have yielded results very similar to humans (Sarmiento, 1975; Harwerth et al., 1995; Siderov and Harwerth, 1995; Prince et al., 2000), this assumption seems reasonable.

It is intriguing to compare the disparity structure of V1 receptive fields with the way in which other parts of the visual system are organized for luminance processing. The response appears to be determined by the disparity presented over a small region, regardless of the disparities outside that area. In this way, their processing of disparity is less sophisticated than the luminance processing of retinal ganglion cells. The closest analogy seems to be with photoreceptors whose response to luminance shows the same simple spatial structure as the response of V1 neurons to disparity. In this sense, disparity selective cells in V1 might be considered the photoreceptors of the “cyclopean retina” (Julesz, 1971).

### 5.2.1 Future work

The properties of V1 neurons do not explain why human sensitivity to depth corrugation is reduced at frequencies lower than 0.3 cpd, as nearly all neurons showed a simple lowpass function of corrugation frequency. The reduced sensitivity to low frequencies may therefore reflect the consequences of processing in extrastriate cortex. Suppose that extrastriate cortex generated a representation of relative disparity by comparing nearby subregions. At very low frequencies, there would be little difference in disparity between the subregions. Consequently the response will be attenuated at low frequencies. Conversely, the lack of low frequency attenuation that I report here for V1 neurons implies that they are not sensitive to disparity differences within their receptive fields. The question arises where sensitivity to disparity differences is first encoded in the brain. Some V2 neurons show selectivity for the disparity within their receptive field relative to their surround (relative disparity, (Thomas et al., 2002). Also, it has been shown that V2

responds selectively to the orientation of disparity defined edges (von der Heydt et al., 2000). V2 therefore seems to be a good candidate to have neurons also specialized for disparity differences within their receptive fields. Examining the spatial properties of these neurons might enable one to identify the neuronal origin of the decline of the human resolution of disparity modulation at low frequencies.

Theoretically, the lower spatial resolution for disparity modulation than for contrast could be a mechanism by which the number of possible matches in the images are reduced, and therefore a step towards solving the correspondence problem (Fleet et al., 1996; Qian and Zhu, 1997). I showed that the low spatial resolution was a direct consequence of the computation of disparity within the energy model. The poor spatial resolution for disparity modulation I found in V1 does therefore not conflict with earlier studies suggesting that these neurons do not solve the correspondence problem (Cumming and Parker, 1997, 2000). The experimental paradigms designed by these authors could be used to test disparity selective neurons in extrastriate cortex for solving the correspondence problem and therefore whether they are more directly linked to perception than V1 neurons.

### 5.3 Temporal frequency tuning for disparity modulation

I compared the psychophysical temporal resolution and the physiological temporal frequency tuning of V1 neurons to disparity modulation. The characteristics of the psychophysical performance are similar to my findings in the population of the V1 neurons: First, the high temporal frequency limitation for both was similar. Second, psychophysical performance was best for values between 1.5 and 3Hz. The averaged peak temporal frequency in the population (2Hz) agreed with this. Third, the psychophysical sensitivity for disparity modulation had a mild attenuation for low temporal frequencies. This was the general pattern in the neurons.

These findings argue in favor of V1 being a limiting temporal factor for the poor temporal resolution of stereopsis. What makes this result surprising is the fact that the temporal response of V1 neurons also seems to match psychophysical detection of

contrast modulation, despite the fact that the psychophysical ability to resolve contrast modulation (de Lange, 1958; Kelly, 1971a, b) is much better than the resolution for disparity modulation (Norcia and Tyler, 1984). This implies that the disparity signal that drives modulation at the input to an individual neuron can change that neuron's temporal response properties. This is confirmed by my comparison of temporal frequency tuning for disparity with temporal frequency tuning for contrast (Figures 4.4, 4.5).

These results can be explained when regarding the computation of disparity in V1 as a cross-correlation of the band-pass filtered monocular inputs. The temporal frequency cutoff for contrast modulation is then determined by the frequency response of the temporal kernel. Conversely, the temporal frequency cutoff for disparity modulation corresponds to the cutoff predicted approximately by the squared width of the monocular bandpass filter.

This also predicts correlations between three different aspects of the response: 1) the rise time of the initial response, 2) the amplitude attenuation with increasing frequency, and 3) the slope of the relationship between phase lag and frequency. I observe exactly such correlations (Figures 4.6B and 4.8). Finally, a recent study (Read and Cumming, 2003c) examined disparity tuning in V1 as a function of interocular delay between the monocular images. The observed integration time for these interocular delays predicted a mean temporal frequency cut off of around 10Hz, closely similar to the mean temporal frequency cut off I found.

These early mechanisms can therefore explain the difference in psychophysical resolution for contrast and for disparity. However, with both types of stimulus, the absolute value of psychophysical resolution is poorer than that of many individual neurons, and indeed poorer than the mean neuronal performance. This discrepancy is not surprising considering that a perceptual decision probably relies on the pooled activity of neurons. Any difference in temporal phase of response modulation between members of the pool (resulting from slightly different temporal integration times) will decrease the modulation

amplitude of the sum of the signals. This phenomenon will have more impact at higher frequencies, shifting the overall temporal resolution downwards

Two other psychophysical phenomena can also be explained by my results. First, when human subjects judge the speed of motion in depth, performance is poor when the only cue is changing disparity in dynamic RDS (Harris and Watamaniuk, 1996). This suggests that the subjects do not have access to explicit neuronal signals about the rate of change of disparity. Examining the phase of neuronal responses to disparity modulation (Figure 4.9) suggests that V1 neurons likewise are not sensitive to the rate of change of disparity over time (motion in depth).

Second, when motion perception is isolated from cues about position, detection thresholds of human observers are lowest for about 2Hz and drop above 5Hz (Nakayama and Tyler, 1981) and above 0.7cpd. These values correspond exactly to those observed for spatial and temporal disparity modulation (Chapters 3 and 4). The energy model (and its functional analogy to cross-correlation) was originally formulated for motion detection (Adelson and Bergen, 1985). If the cross-correlation occurs over the same spatio-temporal filters to construct motion detectors as disparity detectors, this predicts the same spatial and temporal resolution and therefore explains the similarity in the psychophysical behavior.

### 5.3.1 Future work

By using drifting sinusoidal luminance gratings whose disparity modulates sinusoidally in depth, one can control both, the monocular temporal frequency and the temporal frequency of the disparity modulation independently: the above scheme makes predictions about how neurons would respond to such a stimulus. Measuring the responses to this novel stimulus would be an independent test of the scheme I propose.

## 5.4 An example for the analogy between space and time in the brain

In chapter 3, I established that over space, the size of receptive fields of disparity selective neurons can explain why the spatial resolution of stereopsis is so much poorer than that for contrast. (The binocular receptive field size corresponds to the squared envelope of the monocular receptive field, Appendix B.) In chapter 4, I demonstrated that the poor temporal resolution of stereopsis depends on the square of the monocular temporal filter (Appendix C). For both, space and time, the same mechanism explains why single V1 neurons can account for the psychophysical performance in tasks, both, with high (contrast) and low (disparity) resolution: the poorer resolution for disparity modulation than for contrast modulation is explained by binocular cross-correlation of the output of the monocular (spatial or temporal) filters. This explanation holds independently of the dimension (space or time) over which the modulation occurs. Here my data show that the dimensions space and time are treated analogously in the brain. When considering how different temporal and spatial modulations (of contrast or disparity) are perceptually, the analogy appears very counter-intuitive. But the mathematics show that this follows naturally from a well known similarity of temporal-spatial processing: the monocular images are bandpass filtered both, spatially and temporally. Measuring the correlation between bandpass filtered images yields frequency responses of the correlation to lower frequencies than for the bandpass filters.

## 5.5 Encoding disparity in V1: binocular cross-correlation as opposed to global feature matching

Psychophysical studies indicate that the perceptual system for stereopsis is not fooled by false matches in a visual scene but solves the correspondence problem (Cumming and Parker, 1997; Cumming and Parker, 2000). It has been suggested that the lower spatial resolution for disparity modulation is a step towards solving the correspondence problem by decreasing the number of possible matches (Marr and Poggio, 1979; Fleet et al., 1996; Qian and Zhu, 1997) and that the lower temporal resolution results from solving the correspondence problem (Banks, personal communication). I have shown in this thesis

that the lower spatial and temporal resolution for disparity modulation than for contrast modulation follow directly from the computation of cross-correlation by the binocular neuron. The low temporal and spatial resolution for disparity modulation I found in V1 neurons is therefore no evidence that these neurons solve the correspondence problem. Together with previous results (Cumming and Parker, 1997), the findings in this thesis provide further evidence that rather than signaling global matches, V1 neurons respond to local matches by computing a cross-correlation of monocular images.

## 5.6 Conclusion

The most important finding presented in this thesis is that disparity selective V1 neurons, although not directly supporting the perception of depth (Cumming and Parker, 1997, 1999, 2000), impose a limit on both, the spatial acuity and the temporal resolution for detecting depth from disparity. This is the first time that this indirect but pivotal role of disparity selective V1 neurons for stereopsis has been established. It is also the first neuronal and computational characterization of the spatial and temporal resolution for signaling disparity modulation in visual neurons.

Having established that disparity sensitive V1 neurons are selective for a uniform disparity throughout their receptive field, one can ask where in the visual pathway selectivity for more complex disparity features is first encoded. Some V2 neurons show selectivity for the disparity within their receptive field relative to their surround [relative disparity, (Thomas et al., 2002)]. V2 therefore seems to be a good candidate to have neurons also specialized for disparity differences within their receptive fields. Further experiments (Cumming and Parker, 1997, 2000) in extrastriate cortex are necessary to test for disparity selective responses that account for perceptual features of stereopsis requiring more complex processing (e.g. responses to global, not local matches).

I suggest that the lower temporal resolution for disparity modulation than for contrast modulation results from measuring the cross-correlation between the temporally bandpass filtered monocular images.

Binocular cross-correlation of the bandpass filtered monocular images explains that both, the spatial and temporal resolution is lower for disparity modulation than for contrast modulation. The same mechanism therefore accounts for the low resolution for disparity modulation over space and over time. Mathematically, this follows directly from the well know similarity of spatial and temporal bandpass filtering of the monocular images. My results therefore provide another example for a computation (binocular cross-correlation of the monocular images) which occurs analogously over space and time in early visual processing.

Taken together, these findings represent a significant step towards understanding the process by which neurons solve the correspondence problem. Rather than trying to match every feature in each eye, finite regions of the images in each eye (in both space and time) are compared by cross-correlation. Without any additional computation this process itself greatly reduces the number of false matches that need to be eliminated in subsequent processing.

## Appendix A: Reducing the effect of eye-movements on the calculation of relative modulation

Saccadic eye movements result in instantaneous changes of phase of the periodic modulation of the spike train. I used the autocorrelation function of the spike train in an attempt to estimate the periodic modulation in firing in a way that was not sensitive to this effect of eye movements. The autocorrelation function of the spike train describes for all  $t$  how well the spike train at time  $t$  correlates with the spike train at time  $t+\tau$ , i.e. it compares the spike train at  $t$  with that at  $t+\tau$ . For a periodic spike train of period  $\lambda$ , its autocorrelation function is maximal for  $\tau = n\lambda$ , for  $n=0,1,2$  etc.. It thus is a description of the periodic modulation of the spike train, centered around 0, irrespective of the phase of the spike train. Although changes in the phase of a response can affect the extent of correlation, (the modulation amplitude in the autocorrelation function), this is small unless the phase changes are frequent. If a phase-shift occurs at  $t_1$ , then this affects the comparison and consequently the extent of correlation for the segment of the spike train:  $t_1 < t < t_1 + \tau$ , while for  $t < t_1$  and  $t > t_1 + \tau$  it remains unaffected. Hence for small  $\tau$  the effect of occasional phase-shifts is small, and the autocorrelation-function best represents the spike-train cleared from the disruptions caused by saccades. Extracting the  $fI$ -component from the autocorrelation function for small  $\tau$  thus provides an estimate of its value that is relatively unaffected by saccades. At very small values of  $\tau$ , the autocorrelation function reflects the intrinsic statistics of a neuron's firing. I therefore calculated the amplitude of  $fI$  modulation in the autocorrelation function from  $\tau = 250$  ms ( $0.5 \lambda$ ) to  $\tau = 750$  ms ( $1.5 \lambda$ ).

## Appendix B: Relationship between the binocular receptive field size and the spatial frequency cut-off in response to a disparity modulating corrugation

The energy model predicts a relationship between the receptive field size and the the spatial frequency cutoff (defined as the spatial frequency at  $2/3$  peak relative

modulation). Here, I will specify this predicted relationship mathematically. Consider vertically-oriented receptive fields. The response to a planar RDS with disparity  $\delta$  is:

$$D(\delta) = \int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} dx \rho_L(x - \delta, y) \rho_R(x, y) = \int_{-\infty}^{+\infty} dy \rho_{Ly}(y) \rho_{Ry}(y) \int_{-\infty}^{+\infty} dx \rho_{Lx}(x - \delta) \rho_{Rx}(x).$$

And the response to a stimulus containing different disparities at different vertical positions (e.g. a sinusoidal corrugation) is:

$$R = \int_{-\infty}^{+\infty} dy \rho_{Ly}(y) \rho_{Ry}(y) \int_{-\infty}^{+\infty} dx \rho_{Lx}(x - \delta(y)) \rho_{Rx}(x)$$

$$\text{so, } \frac{D(\delta(y))}{\int_{-\infty}^{+\infty} dy' \rho_{Ly}(y') \rho_{Ry}(y')} = \int_{-\infty}^{+\infty} dx \rho_{Lx}(x - \delta(y)) \rho_{Rx}(x).$$

Thus,

$$R = \frac{\int_{-\infty}^{+\infty} dy \rho_{Ly}(y) \rho_{Ry}(y) D(\delta(y))}{\int_{-\infty}^{+\infty} dy' \rho_{Ly}(y') \rho_{Ry}(y')}.$$

Assume that both receptive fields have identical profiles along the  $y$ -axis, Gaussians with standard deviation  $\sigma_y$ . Then the response is:

$$R = \frac{1}{\sigma_y \sqrt{\pi}} \int_{-\infty}^{+\infty} dy e^{-\frac{y^2}{\sigma_y^2}} D(\delta(y)).$$

Note that the responses  $D(\delta(y))$  are convolved with a Gaussian with a standard deviation

$$\frac{1}{\sqrt{2}} \sigma_y, \text{ not } \sigma_y.$$

Now consider the sinusoidal modulation:

$D(\delta(y))$  is a linear function of disparity (approximately) over the range used in the experiments.  $\delta(y)$  is a sinusoidal function of  $y$ .

Thus,

$$R = \frac{1}{\sigma_y \sqrt{\pi}} \int_{-\infty}^{+\infty} dy e^{-\frac{y^2}{\sigma_y^2}} (a \sin(2\pi f y + \phi) + b)$$

This has a constant term plus some modulation. Now the question is how the amplitude of the modulation varies as a function of  $f$ .

The modulating term is

$$M(\phi) = \frac{a}{\sigma_y \sqrt{\pi}} \int_{-\infty}^{+\infty} dy e^{-\frac{y^2}{\sigma_y^2}} \sin(2\pi f y + \phi)$$

(Gradshteyn and Ryzhik, 2000), 3.923.1:  $\int_{-\infty}^{+\infty} dx e^{-ax^2} \sin(2qx + r) = \sqrt{\frac{\pi}{a}} e^{-\frac{q^2}{a}} \sin r$ .

Here,  $a=1/\sigma_y^2$ ,  $q = \pi f$ ,  $r=\phi$ .

$$M(\phi) = a e^{-\sigma_y^2 \pi^2 f^2} \sin \phi.$$

Thus, the amplitude of the modulation is proportional to  $e^{-\sigma_y^2 \pi^2 f^2}$ . This is a Gaussian in frequency with an SD of  $\sigma_{\text{freq}} = 1 / (\sigma_y \pi \sqrt{2})$ .

Now, if the receptive field is probed with a bar of random-dot pattern aligned along the x axis. Assume the image is a delta function in y. The convolution in each eye is

$$v_L(y_{\text{bar}}) = \rho_{Ly}(y_{\text{bar}}) \int dx \rho_{Lx}(x) I_{Lx}(x).$$

On average, therefore, the convolution is proportional to  $\rho_{Ly}(y_{\text{bar}})$ . So, the response of the cell to this monocular bar is

$$C \propto \{\rho_{Ly}(y)\}^2 \propto e^{-\frac{y^2}{\sigma_y^2}}.$$

And the standard deviation one obtains by fitting will be:

$$\sigma_{\text{barprobe}} = \sigma_y / \sqrt{2}.$$

The spatial frequency cutoff in response to the corrugation is the two-thirds maximum point of the Gaussian which is modulation as a function of frequency. Thus, to good accuracy, the corrugation cutoff =  $\sigma_{\text{freq}} = 1 / (\sigma_y \pi \sqrt{2})$ .

This is plotted against  $1 / (2\pi\sigma_{\text{barprobe}}) = \sqrt{2} / (2\pi\sigma_y) = 1 / (\pi\sigma_y \sqrt{2})$ .

For the energy model it should therefore lie on the identity line.

## Appendix C: The effect of temporal filtering of the monocular images on the cross-correlation measured between the images

The energy model can be understood as functionally equivalent to measuring the cross-correlation between monocular images. Suppose that the correlation between the monocular input varies as a sinusoidal function of time (as in the RDS I used whose

disparity varied sinusoidally over time). In the following I will examine the effect of temporal filtering of the monocular images on the frequency response of this cross-correlation. Suppose the images presented to left and right eyes are each a 1-pixel time-series which is a sequence of black and white dots (luminance +/-1). There are no correlations across time; however, the left and right time-series are correlated at any moment in time, and the correlation between left and right images is changing moment to moment: it is a function of time  $c(t)$ . The question now is how the correlation between the filtered time series behaves when the left and right images are filtered:

The images  $I_L(t)$ ,  $I_R(t)$  are passed through temporal filters  $\rho_L$  and  $\rho_R$  to obtain the filtered time-series  $L(t)$ ,  $R(t)$ :

$$L(t) = \int_{-\infty}^t dt' I_L(t') \rho_L(t' - t) \text{ and } R(t) = \int_{-\infty}^t dt' I_R(t') \rho_R(t' - t).$$

Now consider the value of the product of the filtered time-series:

$$LR(t) = \int_{-\infty}^t dt' I_L(t') \rho_L(t' - t) \int_{-\infty}^t dt'' I_R(t'') \rho_R(t'' - t)$$

$$LR(t) = \int_{-\infty}^t dt' \int_{-\infty}^t dt'' \rho_L(t' - t) \rho_R(t'' - t) I_L(t') I_R(t'').$$

The average value over all random images will be written with angle brackets  $\langle \rangle$ :

The original images were uncorrelated from moment to moment. Thus,  $\langle I_L(t') I_R(t'') \rangle = 0$  unless  $t' = t''$ . This gives a Dirac delta (Bracewell, 1986):

$$\langle LR \rangle(t) = \Delta t \int_{-\infty}^t dt' \rho_L(t' - t) \rho_R(t' - t) \langle I_L(t') I_R(t') \rangle$$

where  $\Delta t$  is the extent of one time-frame (temporal pixel).

The value of  $\langle I_L(t') I_R(t') \rangle$  depends on the instantaneous correlation between the images.

This is  $c(t')$ . So,

$$\langle LR \rangle(t) = \Delta t \int_{-\infty}^t dt' \rho_L(t' - t) \rho_R(t' - t) c(t'),$$

or

$$\langle LR \rangle(t) = \Delta t \int_{-\infty}^t dt' \rho^2(t' - t) c(t')$$

if the temporal kernels of the left and right receptive fields are identical.

This is the correlation pattern convolved with the product of the left and right temporal receptive fields. Thus, the Fourier transform of  $\langle LR \rangle$  is the product of the Fourier transform of  $(\rho_L \rho_R)$  or  $\rho^2$  and the Fourier transform of the correlation pattern.

If the correlation pattern is a sinewave,  $c(t')$ , then  $\langle LR \rangle$  will be a sinewave too. The amplitude of the sinewave will be the amplitude of the Fourier transform of  $(\rho_L \rho_R)$  or  $\rho^2$  at that frequency. Suppose that  $\rho$  is a temporal bandpass filter. The squared filter ( $\rho^2$ ) then approximates the squared envelope of the filter, regardless of the temporal structure within it. It therefore approximates a low-pass filter. As a consequence the monocular bandpass filters give rise to a lowpass response for correlation changes.

# Bibliography

- Adelson EH, Bergen JR (1985) Spatiotemporal energy models for the perception of motion. *J Opt Soc Am A* 2:284-299.
- Albrecht DG (1995) Visual cortex neurons in monkey and cat: effect of contrast on the spatial and temporal phase transfer functions. *Vis Neurosci* 12:1191-1210.
- Albright TD (1993) Cortical processing of visual motion. In: *Visual Motion and Its Role in the Stabilization of Gaze* (Miles FA, Wallmann J, eds), pp 177-201. Amsterdam: Elsevier Science Publishers.
- Allman J, Miezin F, McGuinness E (1985a) Stimulus specific responses from beyond the classical receptive field: neurophysiological mechanisms for local-global comparisons in visual neurons. *Ann Rev Neurosci* 8: 407-430.
- Allman J, Miezin F, McGuinness E (1985b) Direction- and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT). *Perception* 14: 105-126.
- Anderson BL, Nakayama K (1994) Toward a general theory of stereopsis: binocular matching, occluding contours, and fusion. *Psychol Rev* 101:414-445.
- Andrews BW, Pollen DA (1979) Relationship between spatial frequency selectivity and receptive field profile of simple cells. *J Physiol* 287:163-176.
- Anzai A, Ohzawa I, Freeman RD (1999a) Neural mechanisms for processing binocular information II. Complex cells. *J Neurophysiol* 82:909-924.
- Anzai A, Ohzawa I, Freeman RD (1999b) Neural mechanisms for processing binocular information I. Simple cells. *J Neurophysiol* 82:891-908.

- Anzai A, Ohzawa I, Freeman RD (2001) Joint-encoding of motion and depth by visual cortical neurons: neural basis of the Pulfrich effect. *Nat Neurosci* 4:513-518.
- Backus BT, Banks MS, van Ee R, Crowell JA (1999) Horizontal and vertical disparity, eye position, and stereoscopic slant perception. *Vision Res* 39:1143-1170.
- Bair W, Cavanaugh JR, Movshon JA (2003) Time course and time-distance relationships for surround suppression in macaque V1 neurons. *J Neurosci* 23: 7690-7701.
- Banks MS, Gepshtein S, Landy MS (2004a) Why is spatial stereoresolution so low? *J Neurosci* 24:2077-2089.
- Banks MS, Gepshtein S, Rose HF (2004b) Do we perceive stereoscopic depth from patches of constant disparity? In: *Vision Science Society Annual Meeting*. Sarasota.
- Barlow HB, Blakemore C, Pettigrew JD (1967) The neural mechanism of binocular depth discrimination. *J Physiol* 193:327-342.
- Berends EM, Erkelens CJ (2001) Strength of depth effects induced by three types of vertical disparity. *Vision Res* 41:37-45.
- Berends EM, van Ee R, Erkelens CJ (2002) Vertical disparity can alter perceived direction. *Perception* 31:1323-1333.
- Beverley KI, Regan D (1974) Visual sensitivity to disparity pulses: evidence for directional selectivity. *Vision Res* 14:357-361.
- Blakemore C, Fiorentini A, Maffei L (1972) A second neural mechanism of binocular depth discrimination. *J Physiol* 226:725-749.

- Blasdel G, Campbell D (2001) Functional retinotopy of monkey visual cortex. *J Neurosci* 21:8286-8301.
- Blasdel GG, Lund JS (1983) Termination of afferent axons in macaque striate cortex. *J Neurosci* 3:1389-1413.
- Blasdel GG, Salama G (1986) Voltage-sensitive dyes reveal a modular organization in monkey striate cortex. *Nature* 321:579-585.
- Boussaoud D, Ungerleider LG, Desimone R (1990) Pathways for motion analysis: cortical connections of the medial superior temporal and fundus of the superior temporal visual areas in the macaque. *J Comp Neurol* 296:462-495.
- Bracewell RN (1986) *The Fourier transform and its applications*, 2nd Edition. Singapore: McGraw-Hill Book Company.
- Bradshaw MF, Rogers BJ (1999) Sensitivity to horizontal and vertical corrugations defined by binocular disparity. *Vision Res* 39:3049-3056.
- Brenner E, Smeets JB, Landy MS (2001) How vertical disparities assist judgements of distance. *Vision Res* 41:3455-3465.
- Brewer AA, Press WA, Logothetis NK, Wandell BA (2002) Visual areas in macaque cortex measured using functional magnetic resonance imaging. *J Neurosci* 22:10416-10426.
- Bridge H, Cumming BG (2001) Responses of macaque V1 neurons to binocular orientation differences. *J Neurosci* 21:7293-7302.
- Bridge H, Cumming BG, Parker AJ (2001) Modeling V1 neuronal responses to orientation disparity. *Vis Neurosci* 18:879-891.

- Brodmann K (1909) Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien, dargestellt auf Grund des Zellenbaues. Leipzig: Barth.
- Burkhalter A, Felleman DJ, Newsome WT, Van Essen DC (1986) Anatomical and physiological asymmetries related to visual areas V3 and VP in macaque extrastriate cortex. *Vision Res* 26:63-80.
- Cagenello R, Rogers BJ (1993) Anisotropies in the perception of stereoscopic surfaces: the role of orientation disparity. *Vision Res* 33:2189-2201.
- Cai D, DeAngelis GC, Freeman RD (1997) Spatiotemporal receptive field organization in the lateral geniculate nucleus of cats and kittens. *J Neurophysiol* 78:1045-1061.
- Campbell FW, Robson JG (1968) Application of Fourier analysis to the visibility of gratings. *J Physiol* 197:551-566.
- Carandini M, Heeger DJ, Movshon JA (1997) Linearity and normalization in simple cells of the macaque primary visual cortex. *J Neurosci* 17:8621-8644.
- Carandini M, Heeger DJ, Movshon JA (1999) Linearity and gain control in V1 simple cell. In: *Cerebral Cortex*. Ulinski PS, Jones EG, Peters A (eds). New York: Plenum.
- Casagrande VA, Norton TT (1991) Lateral geniculate nucleus: a review of its physiology and function. In: *Vision and visual dysfunction. The neural basis of visual function*. (Leventhal AG, ed). Boca Raton, FL: CRC Press.
- Cavanaugh JR, Bair W, Movshon JA (2002) Selectivity and spatial distribution of signals from the receptive field surround in macaque V1 neurons. *J Neurophysiol* 88:2547-2556.

- Cogan AI, Lomakin AJ, Rossi AF (1993) Depth in anticorrelated stereograms: effects of spatial density and interocular delay. *Vision Res* 33:1959-1975.
- Colby CL, Gattass R, Olson CR, Gross CG (1988) Topographical organization of cortical afferents to extrastriate visual area PO in the macaque: a dual tracer study. *J Comp Neurol* 269:392-413.
- Cumming BG (2002) An unexpected specialization for horizontal disparity in primate primary visual cortex. *Nature* 418:633-636.
- Cumming BG, Judge SJ (1986) Disparity-induced and blur-induced convergence eye movement and accommodation in the monkey. *J Neurophysiol* 55:896-914.
- Cumming BG, Parker AJ (1997) Responses of primary visual cortical neurons to binocular disparity without depth perception. *Nature* 389:280-283.
- Cumming BG, Parker AJ (1999) Binocular neurons in V1 of awake monkeys are selective for absolute, not relative, disparity. *J Neurosci* 19:5602-5618.
- Cumming BG, Parker AJ (2000) Local disparity not perceived depth is signaled by binocular neurons in cortical area V1 of the Macaque. *J Neurosci* 20:4758-4767.
- Cumming BG, DeAngelis GC (2001) The physiology of stereopsis. *Annu Rev Neurosci* 24:203-238.
- Daniel PM, Whitteridge D (1961) The representation of the visual field on the cerebral cortex in monkeys. *J Physiol (Paris)* 159:203-221.
- Davison AC, Hinkley DV (1997) *Bootstrap methods and their application*: Cambridge University Press.

de Lange H (1958) Research into the dynamic nature of the human fovea-cortex systems with intermittent and modulated light. II. phase shift in brightness and delay in color perception. *J Opt Soc Am A* 48:784-789.

De Valois RL, Albrecht DG, Thorell LG (1982) Spatial frequency selectivity of cells in macaque visual cortex. *Vision Res* 22:545-559.

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

DeAngelis GC (2000) Seeing in three dimensions: the neurophysiology of stereopsis. *Trends Cogn Sci* 4:80-90.

DeAngelis GC, Ohzawa I, Freeman RD (1991) Depth is encoded in the visual cortex by a specialized receptive field structure. *Nature* 352:156-159.

Dow BM, Snyder AZ, Vautin RG, Bauer R (1981) Magnification factor and receptive field size in foveal striate cortex of the monkey. *Exp Brain Res* 44:213-228.

Dubowitz DJ, Chen DY, Atkinson DJ, Scadeng M, Martinez A, Andersen MB, Andersen RA, Bradley WJ (2001) Direct comparison of visual cortex activation in human and non-human primates using functional magnetic resonance imaging. *J Neurosci Methods* 107:71-80.

Enroth-Cugell C, Robson JG (1966) The contrast sensitivity of retinal ganglion cells of the cat. *J Physiol* 187:517-552.

Enroth-Cugell C, Robson JG (1984) Functional characteristics and diversity of cat retinal ganglion cells. Basic characteristics and quantitative description. *Invest Ophthalmol Vis Sci* 25:250-267.

- Erkelens CJ, Collewijn H (1985) Motion perception during dichoptic viewing of moving random-dot stereograms. *Vision Res* 25:583-588.
- Felleman DJ, Van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* 1:1-47.
- Felleman DJ, Burkhalter A, Van Essen DC (1987) Visual area PIP: an extrastriate cortical area in the posterior intra parietal sulcus of macaque monkeys. In: *Soc Neurosci Abstr*, p 626.
- Fitzpatrick V, Pasnak R, Tyler ZE (1982) The effect of familiar size at familiar distances. *Perception* 11:85-91.
- Fleet DJ, Wagner H, Heeger DJ (1996) Neural encoding of binocular disparity: energy models, position shifts and phase shifts. *Vision Res* 36:1839-1857.
- Foster KH, Gaska JP, Nagler M, Pollen DA (1985) Spatial and temporal frequency selectivity of neurones in visual cortical areas V1 and V2 of the macaque monkey. *J Physiol* 365:331-363.
- Freeman RD, Ohzawa I, Walker G (2001) Beyond the classical receptive field in the visual cortex. *Prog Brain Res* 134:157-170.
- Gardner JL, Anzai A, Ohzawa I, Freeman RD (1999) Linear and nonlinear contributions to orientation tuning of simple cells in the cat's striate cortex. *Vis Neurosci* 16:1115-1121.
- Gibson JJ (1950) *The perception of the visual world*. Boston: Houghton Mifflin.
- Gradshteyn IS, Ryzhik IM (2000) *Table of Integrals, Series, and Products*, 6th Edition. San Diego, USA: Academic Press.

- Grinvald A, Lieke E, Frostig RD, Gilbert CD, Wiesel TN (1986) Functional architecture of cortex revealed by optical imaging of intrinsic signals. *Nature* 324:361-364.
- Guld C, Bertulis A (1976) Representation of fovea in the striate cortex of vervet monkey, *Cercopithecus aethiops pygerythrus*. *Vision Res* 16:629-631.
- Harris JM, Watamaniuk SN (1996) Poor speed discrimination suggests that there is no specialized speed mechanism for cyclopean motion. *Vision Res* 36:2149-2157.
- Harline HK (1938) The response of single optic nerve fibers of the vertebrate eye to illumination on the retina. *Am J Physiol* 121:400-415.
- Hartline HK (1940) The receptive fields of optic nerve fibers. *Am J Physiol* 130:690-699.
- Harwerth RS, Smith EL, 3rd, Siderov J (1995) Behavioral studies of local stereopsis and disparity vergence in monkeys. *Vision Res* 35:1755-1770.
- Hawken MJ, Shapley RM, Grosf DH (1996) Temporal-frequency selectivity in monkey visual cortex. *Vis Neurosci* 13:477-492.
- Heeger DJ (1991) Non-linear model of neural responses in cat visual cortex. In: *Computational Models of Visual Processing* (Landy MS, Movshon JA, eds). Massachusetts: MIT Press.
- Heeger DJ (1993) Modeling simple-cell direction selectivity with normalized, half-squared, linear operators. *J Neurophysiol* 70:1885-1898.
- Helmholtz, von HLF (1867) *Handbuch der physiologischen Optik*. Leipzig: Voss.

- Hendrickson AE, Wilson JR, Ogren MP (1978) The neuroanatomical organization of pathways between the dorsal lateral geniculate nucleus and visual cortex in Old World and New World primates. *J Comp Neurol* 182:123-136.
- Hinkle DA, Connor CE (2002) Three-dimensional orientation tuning in macaque area V4. *Nat Neurosci* 5:665-670.
- Holub RA, Morton-Gibson M (1981) Response of Visual Cortical Neurons of the cat to moving sinusoidal gratings: response-contrast functions and spatiotemporal interactions. *J Neurophysiol* 46:1244-1259.
- Horn BKP, Brooks MJ (1989) *Shape from shading*. Cambridge, Mass: MIT Press.
- Howard IP, Rogers BJ (1995) *Binocular vision and stereopsis*. Oxford, New York: Oxford University Press, Clarendon Press.
- Hubel DH, Wiesel TN (1968) Receptive fields and functional architecture of monkey striate cortex. *J Physiol* 195:215-243.
- Hubel DH, Wiesel TN (1970) Stereoscopic vision in macaque monkey. Cells sensitive to binocular depth in area 18 of the macaque monkey cortex. *Nature* 225:41-42.
- Hubel DH, Wiesel TN (1972) Laminar and columnar distribution of geniculo-cortical fibers in the macaque monkey. *J Comp Neurol* 146:421-450.
- Hubel DH, Wiesel TN (1974) Uniformity of monkey striate cortex: a parallel relationship between field size, scatter, and magnification factor. *J Comp Neurol* 158:295-305.
- Janssen P, Vogels R, Orban GA (1999) Macaque inferior temporal neurons are selective for disparity-defined three-dimensional shapes. *Proc Natl Acad Sci U S A* 96:8217-8222.

- Janssen P, Vogels R, Orban GA (2000a) Three-dimensional shape coding in inferior temporal cortex. *Neuron* 27:385-397.
- Janssen P, Vogels R, Orban GA (2000b) Selectivity for 3D shape that reveals distinct areas within macaque inferior temporal cortex. *Science* 288:2054-2056.
- Janssen P, Vogels R, Liu Y, Orban GA (2001) Macaque inferior temporal neurons are selective for three-dimensional boundaries and surfaces. *J Neurosci* 21:9419-9429.
- 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.
- Judge SJ, Richmond BJ, Chu FC (1980) Implantation of magnetic search coils for measurement of eye position: an improved method. *Vision Res* 20:535-538.
- Julesz B (1971) *Foundations of Cyclopean Perception*: Univ. Chicago Press.
- Julesz B, Payne RA (1968) Differences between monocular and binocular stroboscopic movement perception. *Vision Res* 8:433-444.
- Keesey UT (1970) Variables determining flicker sensitivity in small fields. *J Opt Soc Am* 60:390-398.
- Kelly DH (1961) Flicker fusion and harmonic analysis. *J Opt Soc Am* 51:917-918.
- Kelly DH (1971a) Theory of flicker and transient responses, II. counterphase gratings. *J Optical Soc America* 61:632-640.

- Kelly DH (1971b) Theory of flicker and transient responses, I. uniform fields. *J Opt Soc Am A* 61:537-546.
- Kelly DH, Boynton RM, Baron WS (1976) Primate flicker sensitivity: psychophysics and electrophysiology. *Science* 194:1077-1079.
- Kulikowski JJ, Bishop PO (1981) Fourier analysis and spatial representation in the visual cortex. *Experientia* 37:160-163.
- Kulikowski JJ, Bishop, PO (1981) Linear analysis of the responses of simple cells in the cat visual cortex. *Exp Brain Res* 44:386-400.
- Kulikowski JJ, Marcelja S, Bishop PO (1982) Theory of spatial position and spatial frequency relations in the receptive fields of simple cells in the visual cortex. *Biol Cybern* 43:187-198.
- Kumar T, Glaser DA (1992) Depth discrimination of a line is improved by adding other nearby lines. *Vision Res* 32:1667-1676.
- Lagarias JC, Reeds JA, Wright MH, Wright PE (1998) Convergence properties of the nelder-mead simplex method in low dimensions. *SIAM J Optimization* 9:112-147.
- LeVay S, Gilbert CD (1976) Laminar patterns of geniculocortical projection in the cat. *Brain Res* 113:1-19.
- Levitt JB, Lund JS (1997) Contrast dependence of contextual effects in primate visual cortex. *Nature* 387:73-76.

Li CY, Li W (1994) Extensive integration field beyond the classical receptive field of cat's striate cortical neurons--classification and tuning properties. *Vision Res* 34:2337-2355.

Linsenmeier RA, Frishman LJ, Jakiela HG, Enroth-Cugell C (1982) Receptive field properties of x and y cells in the cat retina derived from contrast sensitivity measurements. *Vision Res* 22:1173-1183.

Livingstone MS, Hubel DH (1987) Connections between layer 4B of area 17 and the thick cytochrome oxidase stripes of area 18 in the squirrel monkey. *J Neurosci* 7:3371-3377.

Longuet-Higgins, HC (1982) The role of the vertical dimension in stereoscopic vision. *Perception* 11:377-386.

Maffei L, Fiorentini A (1973) The visual cortex as a spatial frequency analyser. *Vision Res* 13:1255-1267.

Maffei L, Fiorentini A (1976) The unresponsive regions of visual cortical receptive fields. *Vision Res* 16:1131-1139.

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

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.

Maunsell JH, van Essen DC (1983) The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *J Neurosci* 3:2563-2586.

- Mayhew JE, Longuet-Higgins HC (1982) A computational model of binocular depth perception. *Nature* 297:376-378.
- Menz MD, Freeman RD (2004) Temporal dynamics of binocular disparity processing in the central visual pathway. *J Neurophysiol* 91:1782-1793.
- Merrill EG, Ainsworth A (1972) Glass-coated platinum-plated tungsten microelectrodes. *Med Biol Eng* 10:662-672.
- Mitchison GJ, McKee SP (1990) Mechanisms underlying the anisotropy of stereoscopic tilt perception. *Vision Res* 30:1781-1791.
- Movshon JA, Thompson ID, Tolhurst DJ (1978a) Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *J Physiol* 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* 283:101-120.
- Nakayama K, Tyler CW (1981) Psychophysical isolation of movement sensitivity by removal of familiar position cues. *Vision Res* 21:427-433.
- Nguyenkim JD, DeAngelis GC (2003) Disparity-based Coding of Three-Dimensional Surface Orientation by Macaque Middle Temporal Neurons. *J Neurosci* 23:7117-7128.

- Nienborg H, Bridge H, Parker AJ, Cumming BG (2004) Receptive field size in V1 neurons limits acuity for perceiving disparity modulation. *J Neurosci* 24:2065-2076.
- Ninio J (1985) Orientational versus horizontal disparity in the stereoscopic appreciation of slant. *Perception* 14:305-314.
- Norcia AM, Tyler CW (1984) Temporal frequency limits for stereoscopic apparent motion processes. *Vision Res* 24:395-401.
- Ohzawa I, Freeman RD (1986a) The binocular organization of simple cells in the cat's visual cortex. *J Neurophysiol* 56:221-242.
- Ohzawa I, Freeman RD (1986b) The binocular organization of complex cells in the cat's visual cortex. *J Neurophysiol* 56:243-259.
- 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 (1997a) Encoding of binocular disparity by complex cells in the cat's visual cortex. *J Neurophysiol* 77:2879-2909.
- Ohzawa I, DeAngelis GC, Freeman RD (1997b) The neural coding of stereoscopic depth. *Neuroreport* 8:iii-xii.
- Orban GA, Kennedy H, Maes H (1981) Response to movement of neurons in areas 17 and 18 of the cat: velocity sensitivity. *J Neurophysiol* 45:1043-1058.

- Orban GA, Kennedy H, Bullier J (1986) Velocity sensitivity and direction selectivity of neurons in areas V1 and V2 of the monkey: influence of eccentricity. *J Neurophysiol* 56:462-480.
- Perkel DJ, Bullier J, Kennedy H (1986) Topography of the afferent connectivity of area 17 in the macaque monkey: a double-labelling study. *J Comp Neurol* 253:374-402.
- Pettigrew JD, Nikara T, Bishop PO (1968) Binocular interaction on single units in cat striate cortex: simultaneous stimulation by single moving slit with receptive fields in correspondence. *Exp Brain Res* 6:391-410.
- Poggio GF, Talbot WH (1981) Mechanisms of static and dynamic stereopsis in foveal cortex of the rhesus monkey. *J Physiol* 315:469-492.
- Poggio GF, Poggio T (1984) The analysis of stereopsis. *Annu Rev Neurosci* 7:379-412.
- Poggio GF, Motter BC, Squatrito S, Trotter Y (1985) Responses of neurons in visual cortex (V1 and V2) of the alert macaque to dynamic random-dot stereograms. *Vision Res* 25:397-406.
- Poggio GF, Gonzalez F, Krause F (1988) Stereoscopic mechanisms in monkey visual cortex: binocular correlation and disparity selectivity. *J Neurosci* 8:4531-4550.
- Poggio GE (1990) Cortical neural mechanisms of stereopsis studied with dynamic random-dot stereograms. *Cold Spring Harb Symp Quant Bio* 55:749-758.
- Poggio GF, Doty RW, Jr., Talbot WH (1977) Foveal striate cortex of behaving monkey: single-neuron responses to square-wave gratings during fixation of gaze. *J Neurophysiol* 40:1369-1391.

- Poggio GF, Baker FH, Mansfield RJ, Sillito A, Grigg P (1975) Spatial and chromatic properties of neurons subserving foveal and parafoveal vision in rhesus monkey. *Brain Res* 100:25-59.
- Prince SJ, Rogers BJ (1998) Sensitivity to disparity corrugations in peripheral vision. *Vision Res* 38:2533-2537.
- Prince SJ, Cumming BG, Parker AJ (2002a) Range and mechanism of encoding of horizontal disparity in macaque V1. *J Neurophysiol* 87:209-221.
- Prince SJ, Pointon AD, Cumming BG, Parker AJ (2000) The precision of single neuron responses in cortical area V1 during stereoscopic depth judgments. *J Neurosci* 20:3387-3400.
- Prince SJ, Pointon AD, Cumming BG, Parker AJ (2002b) Quantitative analysis of the responses of V1 neurons to horizontal disparity in dynamic random-dot stereograms. *J Neurophysiol* 87:191-208.
- Purpura K, Tranchina D, Kaplan E, Shapley RM (1990) Light adaptation in the primate retina: analysis of changes in gain and dynamics of monkey retinal ganglion cells. *Vis Neurosci* 4:75-93.
- Qian N, Zhu Y (1997) Physiological computation of binocular disparity. *Vision Res* 37:1811-1827.
- Read JC, Cumming BG (2003a) Measuring V1 receptive fields despite eye movements in awake monkeys. *J Neurophysiol*.
- Read JC, Cumming BG (2003b) Testing quantitative models of binocular disparity selectivity in primary visual cortex. *J Neurophysiol* 90:2795-2817.

- Read JC, Cumming BG (2003c) The neural basis of the Pulfrich effect in the monkey. Soc Neurosci Abst 33:339.7.
- Read JC, Cumming BG (2004) Understanding the cortical specialization for horizontal disparity. Neural Comput in press.
- Read JC, Parker AJ, Cumming BG (2002) A simple model accounts for the response of disparity-tuned V1 neurons to anticorrelated images. Vis Neurosci 19:735-753.
- Regan D, Beverley KI (1973) Some dynamic features of depth perception. Vision Res 13:2369-2379.
- Reid RC, Victor JD, Shapley RM (1992) Broadband temporal stimuli decrease the integration time of neurons in cat striate cortex. Vis Neurosci 9:39-45.
- Rockland KS, Van Hoesen GW (1994) Direct temporal-occipital feedback connections to striate cortex (V1) in the macaque monkey. Cereb Cortex 4:300-313.
- Rogers B, Graham M (1979) Motion parallax as an independent cue for depth perception. Perception 8:125-134.
- Rogers B, Graham M (1982) Similarities between motion parallax and stereopsis in human depth perception. Vision Res 22:261-270.
- Rogers BJ, Bradshaw MF (1993) Vertical disparities, differential perspective and binocular stereopsis. Nature 361:253-255.
- Sakata H, Taira M, Kusunoki M, Murata A, Tsutsui K, Tanaka Y, Shein WN, Miyashita Y (1999) Neural representation of three-dimensional features of manipulation objects with stereopsis. Exp Brain Res 128:160-169.

- Sarmiento RF (1975) The stereoacuity of macaque monkey. *Vision Res* 15:493-498.
- Schiller PH, Finlay BL, Volman SF (1976a) Quantitative studies of single-cell properties in monkey striate cortex. I. Spatiotemporal organization of receptive fields. *J Neurophysiol* 39:1288-1319.
- Schiller PH, Finlay BL, Volman SF (1976b) Quantitative studies of single-cell properties in monkey striate cortex. III. Spatial frequency. *J Neurophysiol* 39:1334-1351.
- Schiller PH, Finlay BL, Volman SF (1976c) Quantitative studies of single-cell properties in monkey striate cortex. V. Multivariate statistical analyses and models. *J Neurophysiol* 39:1362-1374.
- Schlittgen R, Streitberg BHJ (2001) *Zeitreihenanalyse*. Munich: Oldenbourg.
- Schumer R, Ganz L (1979) Independent stereoscopic channels for different extents of spatial pooling. *Vision Res* 19:1303-1314.
- Sekuler R (1974) Spatial vision. *Annu Rev Psychol* 25:195-232.
- Shapley R, Lennie P (1985) Spatial frequency analysis in the visual system. *Annu Rev Neurosci* 8:547-583.
- Shikata E, Tanaka Y, Nakamura H, Taira M, Sakata H (1996) Selectivity of the parietal visual neurones in 3D orientation of surface of stereoscopic stimuli. *Neuroreport* 7:2389-2394.
- Shipp S, Zeki S (1989) The Organization of Connections between Areas V5 and V1 in Macaque Monkey Visual Cortex. *Eur J Neurosci* 1:309-332.

- Siderov J, Harwerth RS (1995) Stereopsis, spatial frequency and retinal eccentricity. *Vision Res* 35:2329-2337.
- Sillito AM, Grieve KL, Jones HE, Cudeiro J, Davis J (1995) Visual cortical mechanisms detecting focal orientation discontinuities. *Nature* 378:492-496.
- 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.
- Stensaas SS, Eddington DK, Dobbelle WH (1974) The topography and variability of the primary visual cortex in man. *J Neurosurg* 40:747-755.
- Sugihara H, Murakami I, Shenoy KV, Andersen RA, Komatsu H (2002) Response of MSTd neurons to simulated 3D orientation of rotating planes. *J Neurophysiol* 87:273-285.
- Taira M, Tsutsui KI, Jiang M, Yara K, Sakata H (2000) Parietal neurons represent surface orientation from the gradient of binocular disparity. *J Neurophysiol* 83:3140-3146.
- Talbot SA, Marshall WH (1941) Physiological studies on neural mechanisms of visual localization and discrimination. *Am J Ophthalmol* 74:1255-1264.
- Tanaka H, Uka T, Yoshiyama K, Kato M, Fujita I (2001) Processing of shape defined by disparity in monkey inferior temporal cortex. *J Neurophysiol* 85:735-744.
- Thomas OM, Cumming BG, Parker AJ (2002) A specialization for relative disparity in V2. *Nat Neurosci* 5:472-478.

- Tootell RB, Hamilton SL, Switkes E (1988a) Functional anatomy of macaque striate cortex. IV. Contrast and magno-parvo streams. *J Neurosci* 8:1594-1609.
- Tootell RB, Silverman MS, Switkes E, De Valois RL (1982) Deoxyglucose analysis of retinotopic organization in primate striate cortex. *Science* 218:902-904.
- Tootell RB, Switkes E, Silverman MS, Hamilton SL (1988b) Functional anatomy of macaque striate cortex. II. Retinotopic organization. *J Neurosci* 8:1531-1568.
- Tootell RB, Hamilton SL, Silverman MS, Switkes E (1988c) Functional anatomy of macaque striate cortex. I. Ocular dominance, binocular interactions, and baseline conditions. *J Neurosci* 8:1500-1530.
- Tootell RB, Silverman MS, Hamilton SL, De Valois RL, Switkes E (1988d) Functional anatomy of macaque striate cortex. III. Color. *J Neurosci* 8:1569-1593.
- Tootell RB, Silverman MS, Hamilton SL, Switkes E, De Valois RL (1988e) Functional anatomy of macaque striate cortex. V. Spatial frequency. *J Neurosci* 8:1610-1624.
- Tootell RB, Hadjikhani NK, Vanduffel W, Liu AK, Mendola JD, Sereno MI, Dale AM (1998) Functional analysis of primary visual cortex (V1) in humans. *Proc Natl Acad Sci U S A* 95:811-817.
- Ts'o DY, Frostig RD, Lieke EE, Grinvald A (1990) Functional organization of primate visual cortex revealed by high resolution optical imaging. *Science* 249:417-420.
- Tsutsui K, Taira M, Min J, Sakai K (1999) Coding of surface orientation by the gradient of texture and disparity in the monkey caudal intraparietal area. In: *Soc Neurosci Abstr*, p 670.

- Tyler CW (1971) Stereoscopic depth movement: two eyes less sensitive than one. *Science* 174:958-961.
- Tyler CW (1973) Steroscopic vision: cortical limitations and a disparity scaling effect. *Science* 181:276-278.
- Tyler CW (1974) Depth perception in disparity gratings. *Nature* 251:140-142.
- Ungerleider LG, Desimone R (1986a) Projections to the superior temporal sulcus from the central and peripheral field representations of V1 and V2. *J Comp Neurol* 248:147-163.
- Ungerleider LG, Desimone R (1986b) Cortical connections of visual area MT in the macaque. *J Comp Neurol* 248:190-222.
- Uttal WR (1981) *A taxonomy of visual processes*. Hillsdale, N.J.: Erlbaum.
- Van Essen DC, Maunsell JH (1980) Two-dimensional maps of the cerebral cortex. *J Comp Neurol* 191:255-281.
- Van Essen DC, Lewis JW, Drury HA, Hadjikhani N, Tootell RB, Bakircioglu M, Miller MI (2001) Mapping visual cortex in monkeys and humans using surface-based atlases. *Vision Res* 41:1359-1378.
- von der Heydt R (1978) Stereoscopic perception of orientation disparity. *Invest Ophthalmol Vis Sci ARVO Abst* 17:286.
- von der Heydt R, Zhou H, Friedman HS (2000) Representation of stereoscopic edges in monkey visual cortex. *Vision Res* 40:1955-1967.

- von Economo C, Koskinas GN (1925) Die Zytoarchitektonik der Hirnrinde des erwachsenen Menschen. Wien: Springer.
- Walker GA, Ohzawa I, Freeman RD (1999) Asymmetric suppression outside the classical receptive field of the visual cortex. *J Neurosci* 19:10536-10553.
- Walker GA, Ohzawa I, Freeman RD (2000) Suppression outside the classical cortical receptive field. *Vis Neurosci* 17:369-379.
- Walker GA, Ohzawa I, Freeman RD (2002) Disinhibition outside receptive fields in the visual cortex. *J Neurosci* 22:5659-5668.
- Wandell BA (1999) Computational neuroimaging of human visual cortex. *Annu Rev Neurosci* 22:145-173.
- Westheimer G (1979) Cooperative neural processes involved in stereoscopic acuity. *Exp Brain Res* 36:585-597.
- Wheatstone C (1838) Contributions to the physiology of vision I.: On some remarkable, and hitherto unobserved, phenomena of vision. *Phil Trans R Soc B* 13:371-395.
- Zeki S (1990) Colour vision and functional specialisation in the visual cortex. Amsterdam: Elsevier.
- Zeki S, Shipp S (1988) The functional logic of cortical connections. *Nature* 335:311-317.
- Zeki SM (1971) Convergent input from the striate cortex (area 17) to the cortex of the superior temporal sulcus in the rhesus monkey. *Brain Res* 28:338-340.
- Zhu YD, Qian N (1996) Binocular receptive field models, disparity tuning, and characteristic disparity. *Neural Comput* 8:1611-1641.

# Publications

The data in this thesis have partially been presented at the following scientific meetings and published as abstracts:

Bridge H., Nienborg H., Cumming B.G., Parker A.: Spatiotemporal Responses of V1 Neurons to Random Dot Stereograms with Sinusoidal Modulation of Disparity, Abstract 164.16, Society for Neuroscience Meeting 2001, San Diego, USA

Nienborg H., Parker A., Cumming B.G.: Receptive field size in V1 may limit acuity for detecting spatial variation in disparity, Abstract 550.19, Society for Neuroscience Meeting 2003, New Orleans, USA

Nienborg H., Bridge H., Cumming B.G., Parker A.: Temporal resolution for disparity modulation may be limited by speed of response modulation in V1, Abstract T169, Vision Sciences Society 4<sup>th</sup> Annual Meeting 2004, Sarasota, USA

Nienborg H., Bridge H., Cumming B.G., Parker A.: Binocular correlation after temporal filtering explains the poor temporal resolution for disparity modulation in V1, Abstract 368.2, Society for Neuroscience Meeting 2004, San Diego, USA

The results in this thesis have been published as the following articles in peer-reviewed scientific journals:

Nienborg H., Bridge H., Cumming B.G., Parker A.: Receptive Field Size in V1 Neurons Limits Acuity for Perceiving Disparity Modulation. *Journal of Neuroscience* (2004); **24**:2065-76

Nienborg H., Bridge H., Cumming B.G., Parker A.: Neuronal Computation of Disparity in V1 limits Temporal Resolution for Detecting Disparity Modulation. *Journal of Neuroscience* (2005); **25**:10207-19

## **Lebenslauf**

### **Hendrikje Nienborg**

Geboren 06.05.1972 in Freiburg/Breisgau

Adresse:

1801 Clydesdale Pl NW, App. 419  
Washington DC, 20009  
USA

Telefon: +1- 202- 332- 9880

e-mail: hn@lsr.nei.nih.gov

### **Berufliche Tätigkeit**

Seit Juni 2003 Postdoktorandin bei Dr. B.G. Cumming,  
Laboratoy of Sensorimotor Research, National Eye Institute,  
National Institutes of Health, Bethesda, MD, USA

### **Ausbildung**

1982 - 91 Hebel-Gymnasium Lörrach  
1991 - 94 Ausbildung in Holzbildhauerei in Berchtesgaden  
1995 - 96 Medizinstudium an der Universität Louis Pasteur in Strasbourg  
1996 -98 Medizinstudium an der Georg-August-Universität in Göttingen  
April 1998 Physikum  
1998 - 2003 Medizinstudium an der Ludwig-Maximilians-Universität in  
München  
April 1999 1. Staatsexamen  
1999 – 2000 Master of Science in Neuroscience, University of Oxford,  
Masterabschluß  
August 2001 2. Staatsexamen Medizin  
April 2003 3. Staatsexamen Medizin

### **Stipendien**

1999-2000 Jahresauslandsstipendium des DAAD  
1999-2001 Studienstipendium der Heinrich-Böll-Stiftung  
1999-2003 Studienstiftung des Deutschen Volkes