



The Spatial Grain of Motion Perception in Human Peripheral Vision

SUSAN J. GALVIN,* DAVID R. WILLIAMS,† NANCY J. COLETTA‡

Received 30 December 1994; in revised form 2 October 1995

Motion reversal effects (the apparent reversal of the direction of motion of a high frequency sinusoidal grating) have been attributed to aliasing by the cone mosaic [Coletta *et al.* (1990). *Vision Research*, 30, 1631–1648] and postreceptoral layers [Anderson & Hess (1990). *Vision Research*, 30, 1507–1515] in human observers. We present data and a new model which suggest that at least two sampling arrays of different densities affect direction discrimination out to 30° eccentricity. The first sampling layer matches anatomical estimates of the cone density. The second sampling layer is too dense to be the parasol cells alone; midget ganglion cells certainly contribute to this task. This is further evidence that motion perception is not mediated exclusively by the magnocellular stream. Copyright © 1996 Elsevier Science Ltd.

Motion Aliasing Sampling Periphery Ganglion cells

INTRODUCTION

An important issue currently under debate is the degree to which different features of the visual scene, such as colour and motion, are processed independently by the visual system. The idea of parallel pathways grew out of the discovery of different ganglion cell types in mammalian retinas with different stimulus selectivities and speeds of transmission (Lennie, 1980). It has been suggested that this functional segregation continues through cortical processing (Zeki, 1978; Ungerleider & Mishkin, 1982; Livingstone & Hubel, 1988). In this formulation, the midget retinal ganglion cells, which are colour selective and support high spatial resolution but transmit information relatively slowly, form the substrate for form perception and project to a temporal cortical stream. The parasol cells, with poorer spatial resolution but higher transmission rates and temporal sensitivity, would provide the only input to the parietal cortical stream, where motion is processed and the locations of objects are coded.

On the other hand, the idea that cortical specializations reflect a direct continuation of the retinal cell populations has been criticized by Merigan and Maunsell (1993). They argue that the retinal specialization is chiefly one of spatiotemporal selectivity, with tuning to higher spatial

and lower temporal frequencies in the parvocellular pathway. Different cortical areas can draw on both the parvocellular and magnocellular pathways in ways appropriate to their specialization. For example, although area MT receives relatively more input from the magnocellular pathway, the parvocellular pathway can support motion perception of stimuli in its spatiotemporal range (Merigan *et al.*, 1991).

We have used the motion reversal effect (Coletta *et al.*, 1990) to estimate the minimum span of motion detectors in human vision. Our results show that motion detectors in peripheral vision have spans that are too small to be explained by the spatial density of neurons in the magnocellular pathway, implicating the more numerous parvocellular neurons in motion perception.

Coletta *et al.* (1990) measured motion reversals attributable to aliasing by the cone mosaic out to 25°, while Anderson and Hess (1990) described an effect at 40° and beyond which they attributed to aliasing at a postreceptoral site. In both of these studies the results were modelled using a single sampling array. We present a new two-stage model of the motion reversal effect which clarifies the roles of cone sampling and postreceptoral sampling in producing the phenomenon, and which suggests that both receptor and post-receptor sampling densities can be estimated from motion reversal data. Our new measurements provide evidence for both receptor and postreceptor aliasing and reveal how their contributions depend on retinal eccentricity.

METHODS

Apparatus

All stimuli were produced by a common-path polar-

* To whom all correspondence should be addressed at: Department of Psychology, University of Otago, P.O. Box 56, Dunedin, New Zealand.

† Center for Visual Science, University of Rochester, Rochester, NY 14627-0270, U.S.A.

‡ College of Optometry, University of Houston, Houston, TX 77204-6052, U.S.A.

ization interferometer described in detail in Sekiguchi *et al.* (1993). The device produces an interference fringe with the contrast, orientation, spatial frequency, and temporal frequency under computer control.

Eye alignment. A bite-bar was used to stabilize the head. The two beams that formed the fringe on the retina were focused in the pupil plane and positioned at the Stiles–Crawford maximum. The correct axial alignment was achieved when either of the beams disappeared abruptly if the head was moved horizontally or vertically, indicating that the beam was meeting the edge of the pupil as a focused spot. The beams were centred on the Stiles–Crawford maximum by setting the spatial frequency to a high value to separate the two beams in the pupil plane, defocusing the field stop to give two discs of light on the retina, and then translating the head to equalize the intensities of the discs. At 20° eccentricity and beyond, it was helpful to alternate the two beams when setting their intensities to be equal, as it was difficult to make this judgement with continuously presented beams.

Fixation. Eccentric viewing was controlled by fixating an LED. All observations were made on the horizontal meridian of the temporal retina of the right eye. The eye was realigned horizontally for each eccentricity, as different rotations of the eyeball place the pupil in different positions relative to the optical axis of the interferometer.

Observers

Three observers participated. ML is emmetropic; SG's slight myopia was corrected with a small shift in the axial position of the field stop. NC is more myopic, and a -0.75 dioptre spherical lens was positioned between the Maxwellian lens and the eye to enable her to bring the field stop into focus. Observations were made with the right eye. The left eye was patched.

Procedure

Stimuli were 100% contrast vertical gratings drifting to the left or right, or horizontal gratings drifting up and down. Measurements were made at 5, 10, 20, 30 and 40° with circular stimulus fields with diameters 2, 3, 6, 10 and 14° respectively. Field size increased with eccentricity to ensure a large number of fringe cycles across the field, even for the low frequencies required in the periphery. An annulus of incoherent light with an outer diameter of 14° was positioned to exactly surround the stimulus field for field sizes less than 14°. This had been found to improve the visibility of the fringe in previous experiments.

At each eccentricity, each observer was tested with a set of 15–20 spatial frequencies (chosen for each eccentricity based on a few pilot runs). Each set was presented 10 times, in a single block of stimulus presentations, with the order of stimulus presentation randomized within each set. Each observer was tested with 10 blocks, for a total of 100 observations per spatial frequency at each eccentricity. All gratings were drifted

at 4 Hz based on evidence from Coletta *et al.* (1990) that the strongest reversals could be obtained near that temporal frequency. We used a two interval forced choice procedure. For example, on trials with vertical gratings, the observer indicated by a button press whether the predominant motion in the stimulus had been to the right in the first interval and left in the second, or to the left in the first interval and right in the second. No feedback was given. Each trial was initiated by the observer's response to the previous trial. If the observer felt he or she had missed seeing the complete stimulus presentation for some reason, he/she signalled that no response be recorded on that trial, and the stimulus was inserted later in the sequence. The task was quite fatiguing, especially at the larger eccentricities, and the observers were encouraged to take rests whenever needed.

The duration of each stimulus interval was 2 sec. This included 500 msec at both the beginning and end of each interval during which the contrast of the fringe was ramped on and off with a Gaussian envelope. The interval between the two stimulus presentations in each trial was 500 msec.

Vertical gratings were used with observers ML and SG. NC had difficulty seeing the motion of vertical gratings at the first eccentricity used to test her (40°) so her trials were run using horizontal gratings. Observer SG also found vertical gratings difficult to see at large eccentricities, so psychometric functions were also obtained from SG at 30 and 40° using horizontal gratings.

RESULTS

The psychometric functions for observers NC, ML and SG are shown in Fig. 1. The five panels in each of A, B and C show data taken at 5, 10, 20, 30 and 40° on one observer. Each point is the mean percent correct based on 100 trials. The error bars give the standard error across ten sessions.

The results show the same general features seen in the data collected by Coletta *et al.* (1990). These are highlighted in Fig. 2, which shows the data from observer NC at 5° eccentricity. As spatial frequency increases, the direction of motion of the gratings becomes less distinct, and performance gradually falls to chance (50% correct). At this point, which Coletta *et al.* called the *first motion null*, there is no dominant direction of motion. At higher frequencies the dominant direction of motion appears to be opposite to its true direction; this is the motion reversal phenomenon. The first spatial frequency at which performance returns to chance after the biggest region of motion reversal was called the *second motion null* by Coletta *et al.* Since this is not always the second point of chance performance in our data, and because Coletta *et al.* found that the spatial frequency at this null matched anatomical estimates of the sampling frequency of the cone mosaic, we refer to it here as the *cone null*. We confirm this relationship between the cone null and cone spacing. A particular goal of this paper is to identify the anatomical substrate of the *first motion null*.

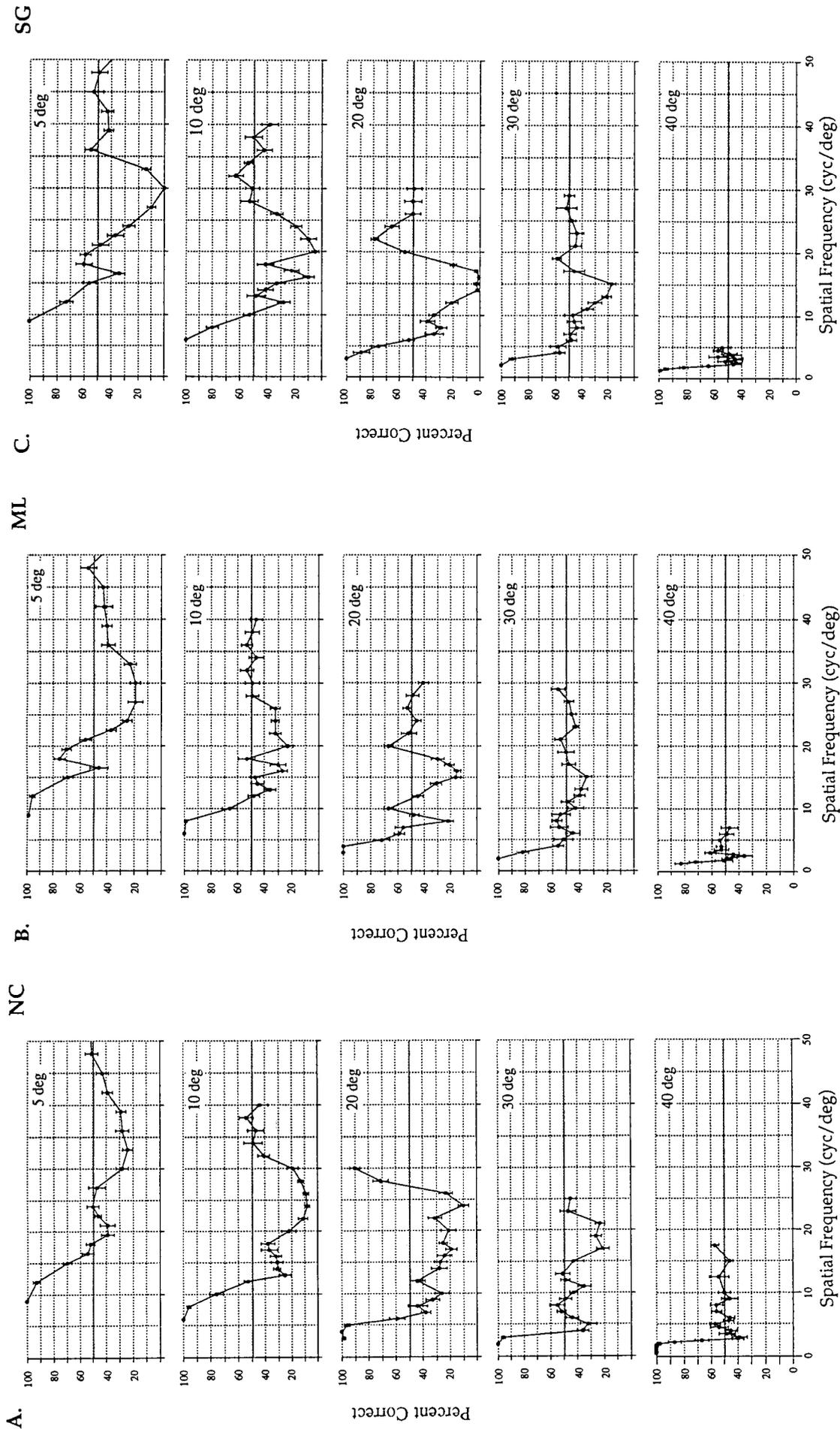


FIGURE 1. Direction discrimination performance for observer NC (A), using horizontal gratings, and for observers ML (B) and SG (C) using vertical gratings. Each point is based on 100 trials, and the error bars show the standard error across ten sessions.

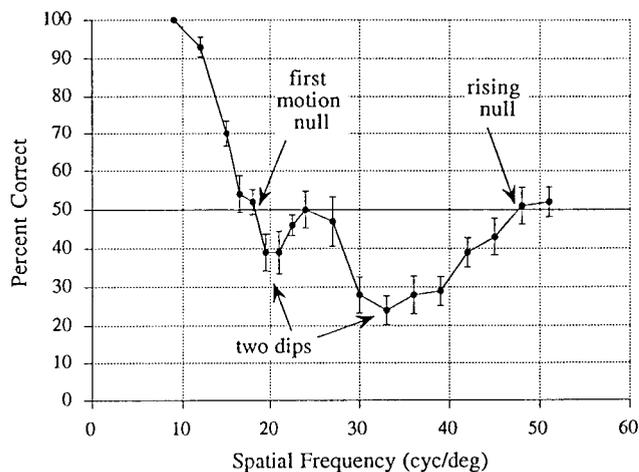


FIGURE 2. Data from observer NC taken at 5° eccentricity, showing the main features of the motion reversal data.

Table 1 lists the first nulls and cone nulls for the three observers. The nulls are located by linear interpolation between the nearest data points.

Motion reversals are seen in the data from all observers at all eccentricities up to 30°. The reversals were strongest (performance usually dropping below 20% correct) within the central 20°, and then became weaker at 30° eccentricity. At 40°, statistically significant reversals were obtained from ML using vertical gratings, and from NC and SG using horizontal gratings. However, in none of these cases did performance fall below 36% correct. Psychometric functions obtained from SG at 40° with vertical gratings did not show reversals; nor did additional functions taken from SG at two other locations slightly above and below the horizontal meridian at this eccentricity.

There are differences in the shapes of the curves taken from different observers at the same eccentricity, and even from the same observer at nearby locations at the same eccentricity. However, at eccentricities out to 30°, almost all the functions reveal two or more dips in the psychometric function between the first motion null and the cone null. A frequently observed pattern of results (see for example Fig. 2) consists of a modest initial motion reversal followed by a more profound motion reversal at higher spatial frequencies. We will argue later that these multiple dips may signal the influence of more than one sampling stage contributing to motion reversals.

INTERMEDIATE DISCUSSION

Comparison with previous studies

The first motion nulls obtained here are consistently lower than those obtained by Coletta *et al.* (1990), particularly beyond 10° eccentricity. The most likely reason for this is that the algorithm they used to locate the first null assumed that their psychometric functions monotonically decreased from the frequency at which performance first fell below 100% to the deepest part of the function. However, the presence of multiple dips can cause the algorithm to overestimate the spatial frequency of the first fall to chance performance. The raw data from the Coletta *et al.* study also showed signs of multiple dips. In our study we sampled more finely in spatial frequency, allowing us to establish the existence of these dips with some certainty, and to determine the position of the first null more accurately.

Coletta *et al.* (1990) presented data taken from observer NC at 10° eccentricity which showed a recovery to better-than-chance performance at spatial frequencies above 30 c/deg. Those data were collected from the nasal retinal meridian using vertical gratings. The data shown in Fig. 1 were taken using horizontal gratings on the temporal retinal meridian.

The only eccentricity at which reversals were found in both this study and that of Anderson and Hess (1990) is 40°. Using vertical gratings, they obtained first nulls at 1.3 and 1.4 c/deg for their two observers, with second motion nulls at 2.6 and 2.8 c/deg, respectively. The only reversal obtained here with a vertical grating was produced by ML, who showed a first null at 1.9 c/deg, followed by a second null at 2.9 c/deg. We obtained a motion reversal from NC using horizontal gratings which showed a first null at 2.6 c/deg. There is clearly variation in the position of the first motion null at 40° between individuals and for different stimulus orientations. The first motion nulls measured here with interference fringes fall within about 50% of those measured by Anderson and Hess, but lie consistently at higher spatial frequencies.

The reversals measured by Anderson and Hess (1990) at 40° were also much deeper than those we measured, with performance falling to 10% in one observer. We tried both horizontal and vertical gratings, but never recorded reversals this strong. Coletta *et al.* (1990) had difficulty finding reversals beyond about 25° with interference fringes, and Artal *et al.* (1995) failed to find

TABLE 1. First nulls and cone nulls for three observers

Eccentricity (deg)	NC (hor)		ML (vert)		SG (vert)	
	First	Cone	First	Cone	First	Cone
5	18.3	47.6	16.3	46.9	15.3	35.3
10	12.1	36.8	11.8	30.5	10.2	27.7
20	6.4	27.2	7.2	19.1	6.2	19.6
30	3.8	23.2	5.3	18.6	5.9	17.6
40	2.6		1.9		2.2	

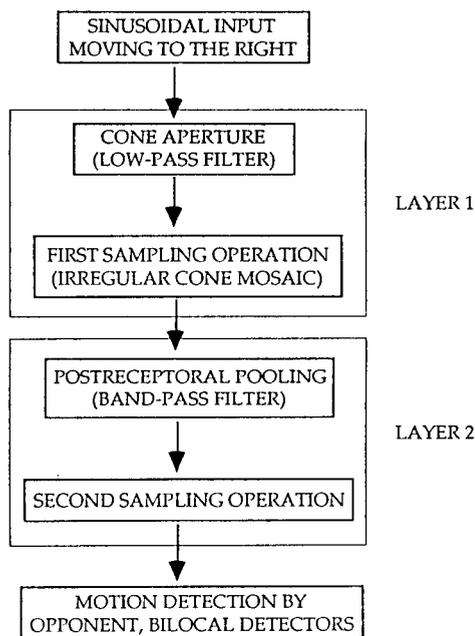


FIGURE 3. Two-stage sampling model.

reversals at 20 and 40° using incoherent light, despite careful refraction. Anderson and Hess instructed their observers to take long (20–30 sec) rests between trials (Anderson, pers. comm.). This may have avoided habituation caused by repeated stimulus presentations in the periphery (Frome *et al.* 1981; Hunzelmann & Spillmann, 1984).

There is a decline in the strength of the motion reversal with increasing eccentricity seen consistently in our data. This could be due to increasing irregularity in the sampling arrays responsible for the effect, as suggested by Artal *et al.* (1995).

MODEL

Anderson and Hess (1990) found first motion nulls that were too low to be accounted for by cone sampling alone, and attributed the reversals they observed at large retinal eccentricities to postreceptoral aliasing, on the grounds that postreceptoral sampling arrays such as the ganglion cells sample more coarsely than the cones in the peripheral retina. Coletta *et al.* (1990), working with interference fringes at smaller eccentricities, concluded that the second motion nulls were due to cone aliasing. They were agnostic about the interpretation of the first motion null because Tiana *et al.* (1991) argued on theoretical grounds that the location of this null could be influenced by spatial pooling as well as sampling.

None of these studies have formally modelled the effects of cascaded sampling stages. Anderson and Hess (1990) remarked that a cone-aliased interference fringe may be further undersampled by postreceptoral processes, but noted that this did not apply to the spatial frequencies at which they were able to measure reversals. Coletta (1992) pointed out that two sampling stages with different Nyquist frequencies would interact in the

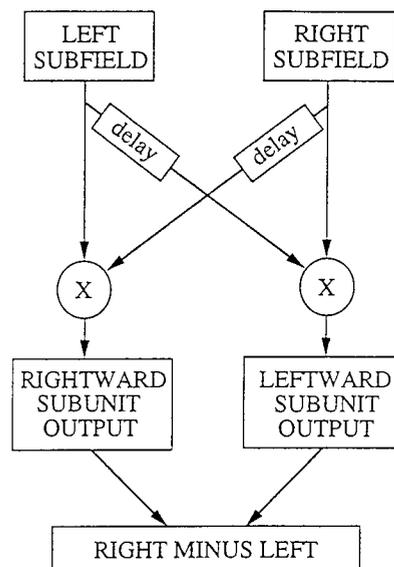


FIGURE 4. Reichardt motion detector. The rightward subunit samples the left subfield then the right subfield and multiplies these values; the leftward subunit samples the right subfield then the left subfield and takes the product of these values. The output of the motion detector is the difference of the two products.

generation of motion nulls. This is because when the stimulus has a frequency higher than the cone Nyquist frequency, the input to the postreceptoral sampling stage is an alias produced by the cone mosaic. The stimulus can therefore be aliased by either, neither, or both the sampling layers, depending on its frequency and the sampling densities of the two arrays. We considered this a possible explanation for the multiple dips in the psychometric functions. Below we show that a model with a single sampling stage does not produce multiple dips, but a model with two stages produces multiple dips which qualitatively resemble our data.

We use a two-stage model to show that the first motion null is probably a good estimate of the Nyquist frequency of the second, coarser sampling array. We also show that a null will occur at a higher frequency matching twice the Nyquist frequency of the first sampling array. We will use these conclusions from our modelling to interpret our empirical data.

The input to the model is a two-dimensional grating with a sinusoidal luminance profile. Figure 3 illustrates the sequence of operations applied to the input. These are described more fully in the Appendix. The sinusoidal input is low-pass filtered to simulate the blurring effect of the cone aperture. Blurring by the optics is not included in the model, as the optics were bypassed experimentally with interference fringes. Next the blurred input is sampled by an irregular sampling array scaled to give it a Nyquist frequency of N_1 . A second sampling array is generated with a lower Nyquist frequency, N_2 , than the first array. Each cell in the second array receives a weighted spatial average of the outputs of nearby first layer cells. The profile of the weighting function is a difference-of-Gaussians function with parameters σ_c and

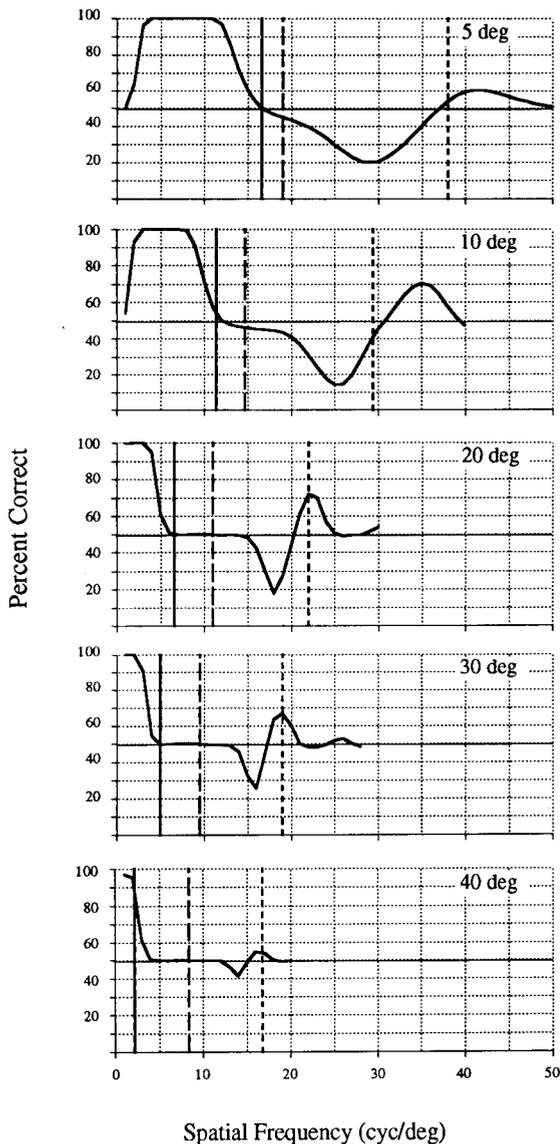


FIGURE 5. Output of single-stage model. The vertical lines are the mean first null of the three observers (solid), and one and two times the cone Nyquist frequency (dashed and dotted lines). The spatial pooling has been made large enough to push the first null down to the mean first null.

σ_s for the excitatory centre mechanism and the inhibitory surround mechanism, respectively. The outputs of every pair of neighbouring second layer cells become the inputs to a layer of bilocal motion detectors. Such detectors are the basis for models of motion perception (Van Santen & Sperling, 1984, 1985; Adelson & Bergen, 1985) that evolved from a model to explain motion reversal in insects (Reichardt, 1961). A rightward-sensitive motion detector is paired with a leftward-sensitive detector, which samples the same subfields but in opposite order. The two detectors become the subunits of an opponent motion detector, seen in Fig. 4. The sign of the difference between their outputs is taken to indicate the direction of motion.

These numbers are summed across all the motion detectors, weighted according to the orientation of the motion detector axis. The weighted sum is converted to

percent correct and plotted against the spatial frequency of the input sinusoid.

The principal differences between this new model and the Tiana *et al.* model are (a) the addition of a second sampling operation, representing a postreceptoral sampling stage, and (b) the calculation of the direction of motion of the sampled stimulus by Reichardt detectors distributed across space, rather than a comparison of the energy in local regions of the spatio-temporal frequency domain. We have used these detectors as they allow us to directly relate the density of the second sampling layer to a simple, plausible model of motion processing. The smallest possible separation between the subfields of the motion detector (the detector's "span") is set by the separation of the adjacent receptive fields of the cells in the pathway that serves motion processing. By estimating the Nyquist frequency of that population, we estimate the minimum motion detector span.

Does the single-stage model account for the data?

Figure 5 shows the output of the model when it contains only a single sampling stage. The two-stage model was reduced to a single-stage model by using identical sampling arrays for both stages. The Nyquist frequency of the single sampling stage was set to anatomical estimates of the cone Nyquist frequency. The five panels show outputs of the single-stage model at the five eccentricities used in the experiment. The leftmost vertical line (solid) shows the mean first null for the three observers. The middle vertical line (dashed) shows the anatomical cone Nyquist frequency and the rightmost vertical line (dotted) indicates twice the cone Nyquist frequency. A single stage of regular sampling with no spatial pooling would give a null at one and two times the cone Nyquist frequency. However, Tiana *et al.* (1991) showed that spatial pooling following sampling can drive the first motion null below the Nyquist frequency of the single sampling stage. We confirm this conclusion, and have added enough spatial pooling at each eccentricity in the model to push the first null of the model down to match the first null in our experimental data. At all eccentricities the first motion null predicted by the model in Fig. 5 simply marks the cutoff of the low-pass filter.

However, this single-stage model requires a very large amount of spatial pooling to explain the observed first motion nulls. Moreover, it does not describe the experimental data well because it does not produce the multiple dips we observed in the data. This suggests that spatial pooling cannot be the complete explanation for the displacement of the first motion null away from the cone Nyquist frequency in the empirical psychometric functions.

Does the two-stage model account for the data?

Two stages of sampling are used to generate the model outputs in Fig. 6. The vertical lines show N_2 and $2N_1$. The parameters N_1 , N_2 , σ_c , and σ_s were adjusted to give nulls that matched the mean first nulls and the cone nulls from

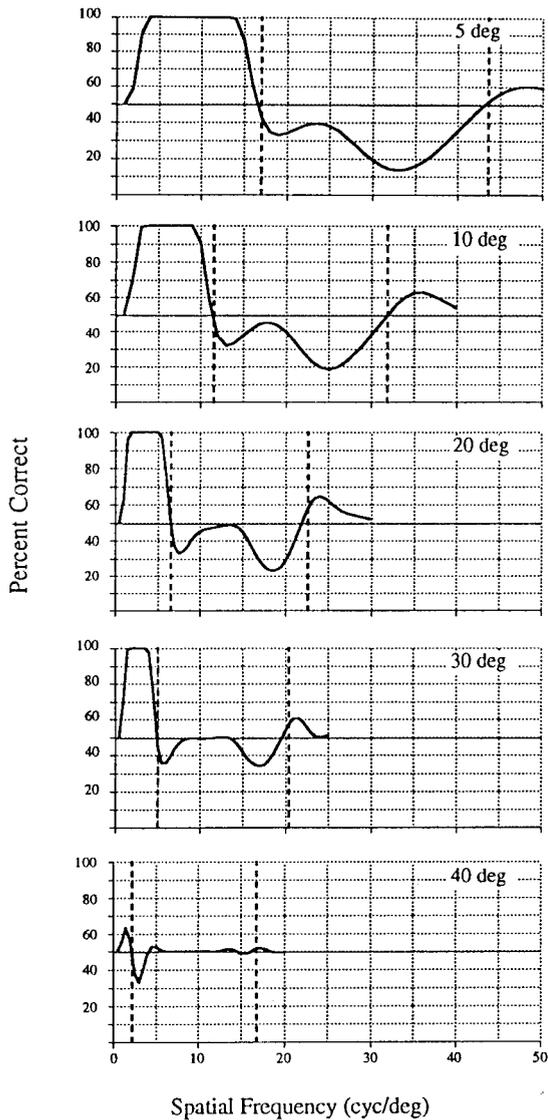


FIGURE 6. Output of the two-stage model. N_2 and $2N_1$ are shown by the vertical lines, and have been chosen so they put the nulls at the mean first nulls and cone nulls for the three observers. Spatial pooling parameters have been chosen so the model outputs reflect the general shape of the empirical curves.

the empirical functions, and to reflect their general shape. The ratio of σ_s to σ_c was set at three. We can see from Fig. 6 that N_2 and $2N_1$ lie very near the nulls they produce at all eccentricities. With the single-stage model, the spatial pools had to be made very large to push the first motion nulls down to the observed values. Here, the match can be achieved with much smaller filters that are in better accordance with other estimates of retinal ganglion cell receptive field profiles (see Galvin, 1994, for details).

Furthermore, note that the two-stage model generates two dips in the psychometric function resembling the multiple dips observed in the experimental data. The important point demonstrated by the two-stage model predictions is that the dip following the first null can be interpreted as a reversal produced by the second, coarser

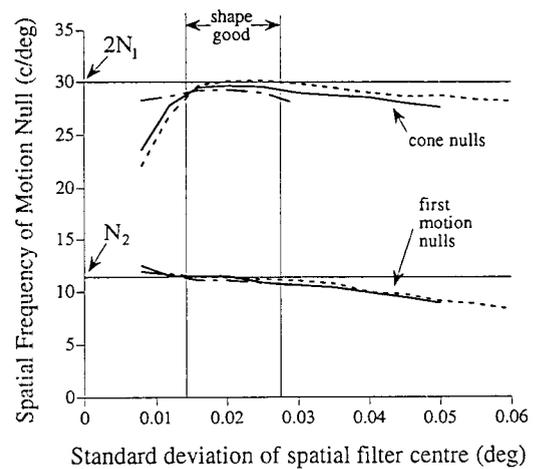


FIGURE 7. Effect of spatial pooling size on the motion nulls. The curves show the first null and cone null positions when the Nyquist frequencies of the first and second layers are fixed and the value of σ_c is varied. The ratio of σ_s to σ_c is 2 for the dotted lines, 3 for the solid lines, and 5 for the broken lines. The vertical lines delimit a range of σ_c which give curve shapes that match the main features of the empirical data.

sampling layer. Though the size and shape of this first dip at higher spatial frequencies can be modified by the sampling effects of the first layer,* the first motion null in the model is almost entirely specified by the Nyquist frequency of the second sampling layer.

Though the first motion null of the two-stage model can be influenced by spatial pooling as it was in the one-stage model, we found that it was a good predictor of the Nyquist frequency of the second sampling layer over a wide range of amounts of pooling. Figure 7 shows how much we would have to increase σ_c and σ_s to affect the position of the first null. The values of N_1 and N_2 have been chosen to be appropriate for the experimental data at 10° eccentricity. The two solid horizontal lines mark the values of N_2 and $2N_1$, and the curves show the first null and the cone nulls produced by the model using different values of σ_c . The three curves near each horizontal line were generated using three different ratios of σ_s to σ_c : three (solid lines), two (dotted lines) and five (broken lines). The vertical lines delimit a range of σ_c values that produce model output shapes similar to the empirical curves. The nulls match N_2 and $2N_1$ very well in this range. Outside this range, the model produces outputs with nulls that are slightly below N_2 and $2N_1$. The values of σ_c below the range of good fits produce wildly fluctuating curves, which could be due to some second layer cells not receiving any input from the first layer mosaic because their receptive fields are too small. When

* One might expect that there should always be a null at the Nyquist frequency of the first sampling layer, since a null produced by the first layer could never be restored to the original signal by some postreceptoral process. Pilot studies with two stages of sampling with regular, one-dimensional arrays showed that a null did occur at N_1 , suggesting that it is the irregularity of the mosaics that prevents performance from returning to chance here in human observers.

σ_c is above the range of good fits, the outputs generated by the model show no dip after the first null, indicating that these are inappropriately large pooling sizes.

We obtained the same results for analyses at 5, 20° and 30°, showing that when the curve exhibits the main features of the empirical data, the first null falls either at N_2 or very slightly below it; the cone nulls never fall more than 5% below or more than 2% above $2N_1$.^{*} The conservative conclusion that can be drawn from this is that the first null and the cone null taken from motion reversal data put *lower bounds* on the Nyquist frequency of the second layer and the sampling frequency of the first layer. The stronger claim is that the nulls can be considered *estimates* of the sampling densities at eccentricities out to 30°, because the nulls produced by the model match N_2 and $2N_1$ in the parameter range that gives shapes roughly like the empirical data.

Though the presence of multiple dips in the data favours a sampling model with more than one sampling stage, we did not thoroughly explore the parameter space to determine which vector of parameter values in the model gave the best fits to the experimental data. There were cases where the values of the parameters we did try in the two-stage model did not fit the data very well. For example, the two-stage model generally predicts two dips in the psychometric function, but not the three dips we have occasionally seen (e.g. observer NC at 30° eccentricity). It is possible that these extra dips reflect more than two sampling stages. If this is the case, we conjecture that the first motion null will still be a measure of the coarsest sampling array in the pathway subserving direction discrimination. When the stimulus matches the Nyquist frequency of the coarsest array, it will be undersampled by that array only, and will appear as a dynamic pattern but without any dominant direction of motion. At frequencies higher than this, but below the Nyquist frequency of the next coarsest array, a dip will be seen. As long as the function does not lie flat along the chance line for frequencies just above the first motion null, we can be confident that it reflects a motion null and not the cutoff frequency of some spatial filter.

There are some aspects of the model which might make the exact shapes of the data hard to fit with any parameter combination, or even make good fits spurious. First, the choice of the function used to transform the motion detector output to percent correct has a big effect on some aspects of the shape of the curves. The Gaussian noise assumed in the model gives a function monotonic with the motion detector output, but affects the steepness of the transitions from one side of the chance line to the other, and the relative depth of different dips. Second, there is the possibility of motion detectors with larger spans affecting the regions of the curve near zero and

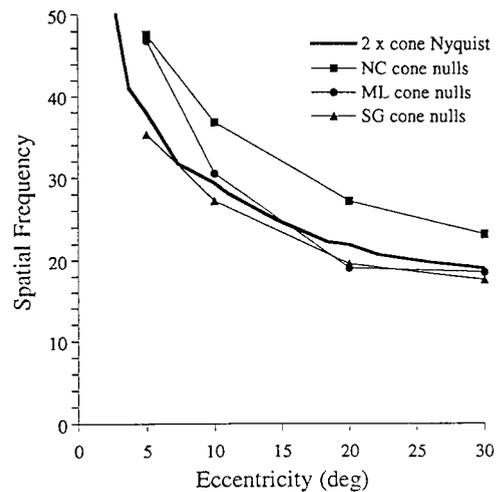


FIGURE 8. Cone nulls from NC (■), ML (●) and SG (▲) compared with $2 \times$ cone Nyquist frequency based on cell counts by Curcio *et al.* (1990). Data were taken using vertical gratings for ML and SG, horizontal gratings for NC.

$2N_1$. Fortunately neither of these things affects the positions of the first null and the cone null. The model captures enough of the shape of the empirical curves to suggest that the dips are caused by second layer aliasing. With that established, we can look to the nulls as the main source of information about the densities of the two sampling functions. We relate these values to possible anatomical substrates in the discussion.

DISCUSSION

Postreceptoral filters do not provide protection from aliasing

Our psychophysical results, in combination with our analysis of multiple-stage sampling, show that the motion reversal effect in the periphery is probably caused by aliasing from at least two sites. This means that postreceptoral spatial pooling is insufficient to protect postreceptoral sampling arrays from aliasing. This is actually a sensible design feature, since modest blurring by the optics and the rarity of high spatial frequencies at high contrast usually makes the total amount of filtering sufficient protection against aliasing under normal viewing conditions (Snyder *et al.*, 1986; Field, 1987; Galvin & Williams, 1992; Williams *et al.*, 1996). If postreceptoral filtering prevented aliasing on its own, this would be a waste of contrast; the postreceptoral filter only needs to supply enough protection to make up for what filters earlier in the system do not provide. Current models of motion detection incorporate protection against aliasing at the levels of the detectors themselves (Van Santen and Sperling, 1984, 1985; Adelson and Bergen, 1985; Watson and Ahumada, 1985; Emerson *et al.*, 1992). The existence of postreceptoral aliasing of interference fringes suggests the neural protection against postreceptoral aliasing is incomplete. The difficulty in observing

^{*} At 40° the ratio of N_2 to N_1 is so small that aliasing by higher multiples of the second layer Nyquist frequency prevents a big cone reversal. There is no null identifiable in the region of twice the cone Nyquist frequency in the model output or the experimental data. The model always produces first nulls at N_2 at 40° eccentricity.

motion reversals with gratings in normal viewing in the near periphery (Artal *et al.*, 1995) suggests that it is the optics, in combination with spatial pooling, that ordinarily prevents motion reversals.

Comparison of motion nulls and anatomical Nyquist frequencies

Figure 8 confirms the finding of Coletta *et al.* (1990) that the cone nulls roughly agree with anatomical estimates of the cone Nyquist frequency. The anatomical data were taken from human temporal retina, based on cell counts by Curcio *et al.* (1990). The conversion from millimetres of retina to degrees of visual angle is taken from Drasdo and Fowler (1974). Cone null values were generally higher for observer NC, who judged horizontal gratings. This is consistent with the finding that cone spacing is greater in the radial direction than in the tangential direction (Curcio & Sloan, 1992), although the effect they found was only a 10–15% difference, compared with a 15–30% difference seen here. The cone mosaic is known to produce spatial aliases under other circumstances (Williams, 1985, 1988, 1992) and is the first population in the visual pathway that has a density which varies with eccentricity in this way, so it is reasonable to conclude that the first layer is the cone mosaic.

The second sampling layer is harder to identify because there are several candidate populations, and their relative densities are hard to determine by anatomical methods. The two largest retinal ganglion cell populations in the human are the midget and parasol ganglion cells. On the basis of dendritic field diameter measurements and the assumption that midget ganglion cell dendritic tree coverage everywhere is two,* Dacey (1993) found that midget ganglion cells make up about 95% of the total in central retina, falling to about 45% in the periphery. Dacey and Petersen (1992) compared the dendritic field sizes of midget and parasol cells, and the degree of coverage by their respective mosaics, and calculated that the ratio of midget to parasol cells is about 30 to 1 at 5° eccentricity and 3 to 1 in the far periphery, making parasol cells about 3% of the total ganglion cell population at 5° and 11% in the periphery.

Figure 9 shows a comparison of anatomical estimates of these ganglion cell Nyquist frequencies and the first motion nulls from the three observers. It is possible that the observers' decisions about the direction of motion were based on the region of the stimulus field nearest the fovea. The open symbols show the nulls plotted at the eccentricities of the inside edge of the fields rather than their centres. The Nyquist frequency estimate for the total ganglion cell population (solid line) is based on counts made by Curcio and Allen (1990). The midget cell Nyquist frequencies (dashed line) are based on measurements of the dendritic field areas of intracellularly

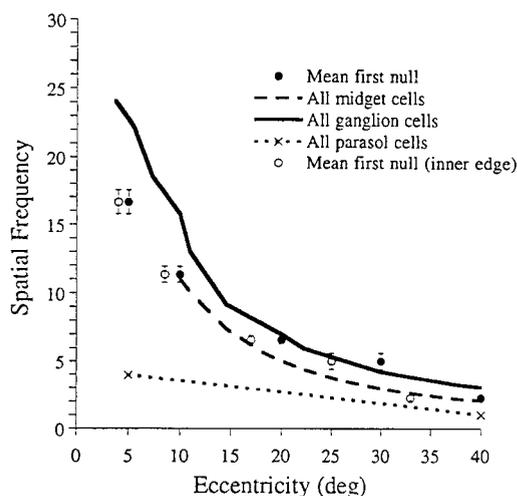


FIGURE 9. First motion nulls and estimates of the Nyquist frequencies of the midget and parasol ganglion cell mosaics based on Dacey (1993) and Dacey and Petersen (1992). The total ganglion cell count comes from temporal retina (Curcio & Allen, 1990). (●) show mean first nulls for three observers, along with standard error; (○) show nulls shifted to the eccentricity of the inside edge of the stimulus field.

injected midget cells from 46 human eyes made by Dacey (1993).† The dotted line shows the trend in parasol density suggested by Dacey and Petersen (1992) for the human retina.

The two-stage model shows that if the first motion null differs from the Nyquist frequency of the second layer, it will fall *below* the Nyquist frequency, so one should look *above* the nulls for possible second layer substrates. It is clear that at all eccentricities the first motion null is much too high to be produced by aliasing by the parasol cell mosaic alone. At 5°, the null is more than four times the parasol Nyquist frequency. If this discrepancy is caused by an underestimate of the parasol density, then it would have to be too low by a factor of more than sixteen. At 40°, where the nulls and parasol densities are most similar, the first null is still twice as high as the estimate of the parasol Nyquist frequency.

The first motion nulls for the three observers follow no single line closely in this figure. The nulls lie a little above the curve for the total midget cell population at 20° and beyond. The dendritic field diameters from Dacey (1993) are pooled over temporal, inferior, and superior retina, and may overestimate the diameters of the cells from temporal retina alone, as the density of the total

* He found a coverage factor of one for each of the on- and off-centre layers.

† Dacey (1993) expresses the size of the dendritic fields at any eccentricity as the diameter of the circle of the same mean area as the cells he measured. We have taken the mean cell spacing to be the diameter of this circle divided by 1.05, which gives the centre-to-centre spacing, s , of an array of hexagons with the same area as this circle. The Nyquist frequency is then equal to $1/\sqrt{3}s$. Dacey's measurements are only used to estimate Nyquist frequencies for 10° eccentricity and beyond because it is likely that the dendritic fields reach a minimum size and begin to overlap at some eccentricity less than this.

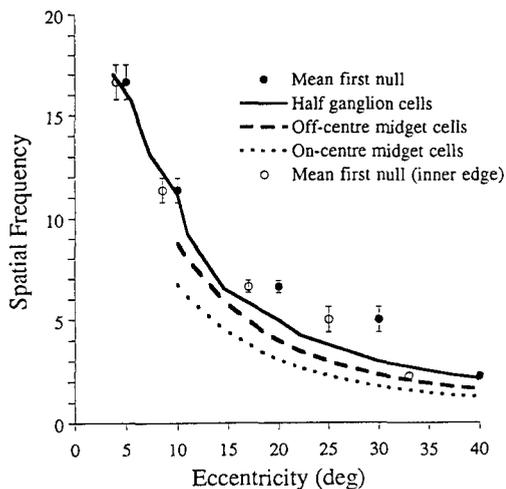


FIGURE 10. Mean of first nulls from the three observers compared with anatomical estimates of the Nyquist frequencies of the on- and off-centre midget ganglion cell populations (dotted and dashed lines respectively), based on dendritic field sizes measured by Dacey (1993). The solid line shows Nyquist frequency for half the total ganglion population.

ganglion cell population in temporal retina is the highest of these three quadrants (Curcio & Allen, 1990). This makes the frequencies plotted here underestimates of the Nyquist frequencies of the midget cell population. This might account for some of the difference between the nulls and the midget Nyquist frequency, but leaves the possibility that the midget cells might not be acting alone in this task.

A complicating factor is the presence of on-centre and off-centre cells making up sub-populations of both midget and parasol ganglion cells. It is not clear whether these should be considered to be sampling the image independently. In the foveal region, where there is an on- and an off-centre midget cell for every cone (Calkins *et al.*, 1994), no advantage would be gained by combining on- and off-centre cells into one sampling array, as it would oversample the cone outputs. However, Dacey (1993) has shown that in the region 25–45° eccentricity, there are 1.7 times as many off-centre midget ganglion cells as there are on-centre cells, and that each subgroup independently tiles the retina. This argues against the idea that each on-centre cell is paired with an off-centre cell in order to encode the whole range of increments and decrements at their shared spatial location. On the other hand, it raises the possibility that each population provides the substrate for an independent representation with its own spatial resolution.

Figure 10 shows the Nyquist frequencies for the on- and off-centre midget ganglion cells, obtained by multiplying the Nyquist frequency of the total midget population by $\sqrt{1/2.7}$ and $\sqrt{1.7/2.7}$, respectively. Also shown is the Nyquist frequency for half of all the ganglion cells. The nearest match seen here is between the shifted nulls and the Nyquist frequency of half the total ganglion cell Nyquist frequency, which is close to

the curve for the total midget ganglion cell Nyquist frequency. It seems that if on- and off-centre ganglion cells do sample the image independently, then the midget cells alone cannot be the substrate for this task.

The data shown here do not allow us to distinguish which selection or combination of on- and off-centre midget and parasol cells make up the substrate for the observed performance. However, they clearly indicate that midget ganglion cells must be involved in this direction discrimination task, and that the parasol cells cannot be the sole substrate for motion perception.

Midget cells contribute to motion processing

The comparison of midget and parasol cell populations with our first motion nulls confirms suggestions that the midget cells must at least contribute to this motion processing task. We have argued that the first motion nulls reflect the Nyquist frequency of a two-dimensional postreceptoral sampling array or, equivalently, one over twice the minimum subfield separation achievable in Reichardt motion detectors built with input at this level of coarseness. Koenderink *et al.* (1985) estimated the minimum span of motion detectors by having observers judge the coherence of random dot fields segmented into strips of alternating, opposite directions of motion. They found spans smaller than a minute of arc for foveal viewing, which also excludes parasol cells as the lone substrate there.

Recently, Anderson *et al.* (1995) came to a similar conclusion about their motion reversal data taken at 40° using natural corrected viewing. Dacey (1993) and Lennie (1993) have noted that estimates of midget ganglion cell Nyquist frequencies match estimates of achromatic acuity made by Anderson *et al.* (1991) across a range of eccentricities. These acuity measures were obtained by extrapolating observers' contrast sensitivity to sinusoidal gratings drifting at 8 Hz. Here again the perception of a moving stimulus is seen to be subject to the same spatial limits as static viewing.

CONCLUSION

We have made two main points in this paper:

1. The motion reversal effect should be modelled with two sampling stages, not just one, as previous studies have used. The two-stage model directs us to the first motion null as an estimate of the Nyquist frequency of the post receptor sampling density, as proposed by Anderson and Hess (1990), and confirms that the first motion null following the big reversal measured out to 30° eccentricity is a good estimate of twice the Nyquist frequency of the cone array, as found by Coletta *et al.* (1990).

2. Comparison of the first motion nulls with anatomical estimates of retinal cell populations shows the sampling density of the post receptor cells underlying the motion reversal effect to be too high to be the parasol cells alone. This adds to accumulating evidence that the magnocellular pathway is not the sole contributor to cortical motion processing.

REFERENCES

- Adelson, E. H. & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America*, 2, 284–299.
- Albright, T. D. (1984). Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of Neurophysiology*, 52, 1106–1130.
- Anderson, S. J., Drasdo, N. & Thompson, C. M. (1995). Parvocellular neurons limit motion acuity in human peripheral vision. *Proceedings of the Royal Society of London, B*, 261, 129–138.
- Anderson, S. J. & Hess, R. F. (1990). Post-receptor undersampling in normal human peripheral vision. *Vision Research*, 30, 1507–1515.
- Anderson, S. J., Mullen, K. T. & Hess, R. F. (1991). Human peripheral spatial resolution for achromatic and chromatic stimuli: limits imposed by optical and retinal factors. *Journal of Physiology*, 442, 47–64.
- Artal, P., Derrington, A. M. & Colombo, E. (1995). Refraction, aliasing, and absence of motion reversals in peripheral vision. *Vision Research*, 35, 939–947.
- Calkins, D., Schein, S. J., Tsukamoto, Y. & Sterling, P. (1994). Ganglion cell circuits in primate fovea. In Drum, B. (Ed.), *Proceedings, colour vision deficiencies, XII* (pp. 267–274). Dordrecht: Kluwer.
- Coletta, N. J. (1992). Effect of luminance on the perceived direction of motion. *Supplement to Investigative Ophthalmology and Visual Science*, 33, 1138.
- Coletta, N. J., Williams, D. R. & Tiana, C. L. M. (1990). Consequences of spatial sampling for human motion perception. *Vision Research*, 30, 1631–1648.
- Curcio, C. A. & Allen, K. A. (1990). Topography of ganglion cells in human retina. *Journal of Comparative Neurology*, 300, 5–25.
- Curcio, C. A. & Sloan, K. R. (1992). Packing geometry of human cone photoreceptors: Variation with eccentricity and evidence for local anisotropy. *Visual Neuroscience*, 9, 169–180.
- Curcio, C. A., Sloan, K. R., Kalina, R. E. & Hendrickson, A. E. (1990). Human photoreceptor topography. *Journal of Comparative Neurology*, 292, 497–523.
- Dacey, D. M. (1993). The mosaic of midget ganglion cells in the human retina. *Journal of Neuroscience*, 13, 5334–5355.
- Dacey, D. M. & Petersen, M. R. (1992). Dendritic field size and morphology of midget and parasol ganglion cells of the human retina. *Proceedings of the National Academy of Sciences*, 89, 9666–9670.
- Derrington, A. M. & Lennie, P. (1984). Spatial and temporal contrast sensitivities of neurones in lateral geniculate nucleus of macaque. *Journal of Physiology*, 357, 219–240.
- Drasdo, N. & Fowler, C. W. (1974). Non-linear projection of the retinal image in a wide angle schematic eye. *British Journal of Ophthalmology*, 58, 709–714.
- Emerson, R. C., Bergen, J. R. & Adelson, E. H. (1992). Directionally selective complex cells and the computation of motion energy in cat visual cortex. *Vision Research*, 32, 203–218.
- Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America*, A, 4, 2379–2393.
- Frome, F. S., MacLeod, D. I. A., Buck, S. L. & Williams, D. R. (1981). Large loss of visual sensitivity to flashed peripheral targets. *Vision Research*, 21, 1323–1328.
- Galvin, S. J. (1994). The spatial grain of motion perception in human peripheral vision. Unpublished Ph.D. thesis, University of Rochester, Rochester, New York.
- Galvin, S. J. & Williams, D. R. (1992). No aliasing at edges in normal viewing. *Vision Research*, 32, 2251–2259.
- Hunzelmann, N. & Spillmann, L. (1984). Movement adaptation in the peripheral retina. *Vision Research*, 24, 1765–1769.
- Koenderink, J. J., van Doorn, A. J. & van de Grind, W. A. (1985). Spatial and temporal parameters of motion detection in the peripheral visual field. *Journal of the Optical Society of America*, A, 2, 252–259.
- Lennie, P. (1980). Parallel visual pathways: a review. *Vision Research*, 20, 561–594.
- Lennie, P. (1993). Roles of M and P pathways. In Shapley, R. & Lam, D. M.-K., (Eds), *Contrast sensitivity* (pp. 201–213). Cambridge, MA: MIT Press.
- Livingstone, M. & Hubel, D. H. (1988). Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science*, 240, 740–749.
- MacLeod, D. I. A., Williams, D. R. & Makous, W. (1992). A visual nonlinearity fed by single cones. *Vision Research*, 32, 347–363.
- Merigan, W. H., Byrne, C. E. & Maunsell, J. H. R. (1991). Does primate motion perception depend on the magnocellular pathway? *Journal of Neuroscience*, 11, 3422–3429.
- Merigan, W. H. & Katz, L. M. (1990). Spatial resolution across the macaque retina. *Vision Research*, 30, 985–991.
- Merigan, W. H. & Maunsell, J. H. R. (1993). How parallel are the primate visual pathways? *Annual Review of Neuroscience*, 16, 369–402.
- Raymond, J. E. (1993). Movement direction analysers: Independence and bandwidth. *Vision Research*, 33, 767–775.
- Reichardt, W. (1961). Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In Rosenblith, W. A. (Ed.), *Sensory communication*. New York: Wiley.
- Sekiguchi, N., Williams, D. R. & Brainard, D. H. (1993). Aberration-free measurements of the visibility of isoluminant gratings. *Journal of the Optical Society of America*, 10, 2105–2117.
- Snyder, A. W., Bossomaier, T. R. J. & Hughes, A. (1986). Optical image quality and the cone mosaic. *Science*, 231, 499–501.
- Tiana, C. L. M., Williams, D. R., Coletta, N. J. & Haake, P. W. (1991). A model of extrafoveal aliasing in human vision. In Landy, M. & Movshon, A. (Eds), *Computational models of visual processing*. Cambridge, MA: MIT Press.
- Ungerleider, L. G. & Mishkin, M. (1982). Two cortical visual systems. In Ingle, D. J., Mansfield, R. J. W. & Goodale, M. S. (Eds), *The analysis of visual behaviour*. Cambridge, MA: MIT Press.
- Van Santen, J. P. H. & Sperling, G. (1984). Temporal covariance model of human motion perception. *Journal of the Optical Society of America*, 1, 451–473.
- Van Santen, J. P. H. & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America*, 2, 300–320.
- Watson, A. B. & Ahumada, A. J. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America*, 2, 322–342.
- Williams, D. R. (1985). Aliasing in human foveal vision. *Vision Research*, 25, 195–205.
- Williams, D. R. (1988). Topography of the foveal cone mosaic in the living human eye. *Vision Research*, 28, 433–454.
- Williams, D. R. (1992). Photoreceptor sampling and aliasing in human foveal vision. In Moore, D. T. (Ed.), *Tutorials in optics* (pp. 15–28). Washington DC: Optical Society of America.
- Williams, D. R., Artal, P., Navarro, R., McMahon, M. & Brainard, D. H. (1996). Off-axis optical quality and retinal sampling in the human eye. *Vision Research*, 36, 1103–1114.
- Zeki, S. M. (1978). Functional specialization in the visual cortex of the rhesus monkey. *Nature*, 274, 423–428.

Acknowledgements—We thank Nobu Sekiguchi, Bill Haake and Al Russell for their assistance with the interferometer. Martin Lankheet donated many hours of observing and discussion time. This work was submitted in partial fulfillment of the Ph.D. requirements at the Center for Visual Science, at the University of Rochester. It was supported by NIH grants EY01319 and EY04367 to David Williams, and NEI grant R29 09431 to Nancy Coletta. The main results in this paper were presented at the 1993 meeting for the Association for Research in Vision and Ophthalmology in Sarasota, Florida, U.S.A.

APPENDIX

Details of Motion Reversal Model

Input defined

The input to the model is a two-dimensional grey-scale pattern, $g(x,t)$, which varies in time, t , and in one dimension in space, x . The stimulus has a sinusoidal luminance profile in the x direction with spatial frequency f_s , mean luminance m , and amplitude a , and always drifts to the right at temporal frequency f_t :

$$g(x,t) = m + a \sin(2\pi f_s x - 2\pi f_t t).$$

The mean luminance is always set to zero as the responses of the Reichardt detectors are independent of it. The amplitude is set to an arbitrary non-zero value (usually one). The temporal frequency is always set to 4 Hz, the value used in the experiments; the spatial frequency is varied.

Cone aperture filter

An attenuating factor is applied to the input value to represent demodulation by averaging across the cone aperture. This factor is a Gaussian function of spatial frequency, with a different standard deviation at each eccentricity. MacLeod *et al.* (1992) made psychophysical estimates of the cone aperture function in coherent light, from which we derived the standard deviations 33.82, 31.92, 29.49, 28.60 and 27.23, which were used at 5, 10, 20, 30 and 40° eccentricity, respectively.

Generating sampling arrays

The user specifies the Nyquist frequency for each of the two layers, N_1 and N_2 , and the length of one side of a square sampling field. The positions for the elements in both arrays are derived from a digitized micrograph of a region of monkey cone mosaic centred at 3.8° eccentricity, supplied by Hugh Perry. Points for the two layers are copied from opposite corners of the array to prevent correlation between them, and each layer is scaled to the appropriate sampling density. Because the micrograph was taken in the peripheral retina, the mosaic is already irregular, and no further jitter is added. Each second layer cell collects input from first layer cells that lie within a pooling range, in a manner described below. If any edge of the field lies within this range of a second layer cell, that second layer cell is removed. This avoids artifacts that would arise from motion detectors that had missing regions in one or both subfields.

Temporal integration

The output of each first layer cell is a discrete integral of the stimulus values that occurred within a few time-steps before time t . Those values are windowed by an exponential temporal integration function with time constant t_c . The exponential filter reduces the input to less than 1% of its original value at $3t_c$. Varying the temporal integration time within a range of plausible values showed that it only scaled the output when the input was a 4 Hz waveform, so the integration time was set to a very short value (1 msec) to reduce computation time.

Sampling in time

Only one cycle of the input sinusoid is sampled, since sampling one period of the sinusoid gives the same motion detector output as sampling an infinite number of them. The total duration of the stimulus is actually a little longer than a cycle in order to accommodate the delay, d , in the motion detector, and the integration time. The total duration is therefore $1/f_t + d + 3t_c$. We always used inputs drifting at 4 Hz, and the delay was always one sixteenth of a second (see motion detection section), so this total was $0.250 + 0.063 + 0.003 = 0.316$ sec.

Postreceptoral pooling

The second layer values are weighted sums of values from the first layer cells lying within a fixed pooling range. Since there is no evidence for the extraction of motion information from the visual

signal earlier than cortex in primates, it is assumed that the signal used for direction discrimination must be band-pass-filtered by cells with a centre-surround receptive field organization established in the retina. The user specifies standard deviations for two Gaussian functions: σ_c for an excitatory centre mechanism and σ_s for a broader inhibitory surround mechanism ($\sigma_s > \sigma_c$). Each of the two functions is truncated at three times its standard deviation. Cells with antagonistic receptive field arrangements usually do not respond well to uniform fields, so the two Gaussians are weighted so that the integral under each of them is the same. The receptive field profile is then:

$$\frac{1}{\sigma_c^2} \exp\left(\frac{-(x^2 + y^2)}{2\sigma_c^2}\right) - \frac{1}{\sigma_s^2} \exp\left(\frac{-(x^2 + y^2)}{2\sigma_s^2}\right).$$

Derrington and Lennie (1984) measured contrast sensitivity functions of single cells from the lateral geniculate nucleus of the macaque, and derived the parameters σ_c and σ_s from best-fitting difference-of-Gaussian functions. These values varied with the temporal frequency at which the contrast sensitivity function was measured. For a drift frequency of 5.2 Hz, the ratio σ_s/σ_c varied between 1.3 and 9.1 for just six parvocellular cells. A ratio σ_s/σ_c of 3.0 was used unless stated otherwise.

Motion detection

The receptive fields of pairs of second layer cells lying within a certain distance of each other become the subfields of the motion detectors. Each cell forms a separate opponent motion detector with any cell within $1.3(\sqrt{3N_2})^{-1}$, that is, within 1.3 times the average cell separation for the mosaic. This produces an average distance between detector subfields which is equal to the average cell separation for the mosaic.

The Reichardt detector gives its maximum amplitude of response across all spatial frequencies when the delay applied to the signal at one subfield is equal to one quarter of the temporal period of the input signal. In order to reflect the performance of cells tuned to our input signal, we set the delay to one sixteenth of a second. (See Galvin, 1994, pp. 16–17 for an explanation of this tuning.)

The final output of each motion detector is the discrete sum of the differences between the leftward and rightward detectors over one temporal cycle of the input sinusoid, and those sums are summed for all the motion detectors to give the output, O .

Orientation tuning

In the two-dimensional sampling mosaic used in this model, most of the motion detector axes (the line joining the centres of the two subfields) are not lined up with the direction of motion of the input sinusoid, perpendicular to its light and dark bands. This means their effective spans are shorter than the distance between the subfields. Since there is no evidence that the visual system is able to take advantage of this higher, interpolated sampling density under other circumstances, we applied orientation tuning to the individual motion detector outputs. Direction tuning of motion sensitive cells in macaque striate cortex (Albright, 1984) and psychophysical estimates of orientation tuning at threshold in the human (Raymond, 1993) can be described by a Gaussian orientation filter with a bandwidth of $\pm 35^\circ$. When this filter is applied to the motion detector spans, the weighted average span is approximately $(2N_2)^{-1}$, the spacing between rows of a triangular lattice of the same average sampling density.

Transforming motion detector output to a performance measure

The sum of all the motion detector outputs, O , gives a summary of the dominant direction of motion signalled by the detectors: right if O is positive, left (and therefore incorrect) if O is negative. Although the left vs right decision is a binary one, it is also clear from the observers' reports that at most spatial frequencies tested they experienced some uncertainty about the direction of motion of the grating. We captured this in the model by centering a Gaussian distribution, representing noise in the system, on the value O . The proportion of the distribution lying above zero was deemed to be the percent correct predicted by the model at that spatial frequency.

The choice of standard deviation for the noise distribution affects the

amplitude of the resulting psychometric function, but not the spatial frequencies at which the curve crosses the chance line, or the number of dips in the function. Since the output of the model is in arbitrary units, the choice of this standard deviation is arbitrary too, and it was chosen so as to make the variations in the psychometric function appreciable by the reader.