Chapter 6

Pattern noise

IN an attempt to get accurate estimates of yaw velocity using an elaborated Reichardt correlator, the effect of *pattern noise* (deviation of the correlator output resulting from the structure of the visual scene) on the correlator response is investigated. Different sampling methods are tested here and it is found that a circular sampled array of elementary motion detectors (EMDs) reduces pattern noise effectively compared to an array of rectangular or randomly selected EMDs for measuring rotational motion.

6.1 Pattern noise

The errors considered here in this chapter are due to *pattern noise*, which is a term given to the deviation in correlator output that results from the structure of the visual scene. Physiological motion detectors also suffer from random noise, which is due to the variation in its response on repeated presentation of identical stimulus patterns. The random noise experienced by a biological motion detector falls into two categories, namely photon noise and intrinsic noise. The photon noise results from variations in the number of photons absorbed by a photoreceptor in a given unit of time. In addition, the neurons and synapses that comprise the correlator generate intrinsic noise. Studies carried out on the LMCs (Lamina Monopolar Cells) by Laughlin (1994) indicates that photon noise dominates intrinsic noise up to moderate light intensities and at higher light intensities, photon noise equals intrinsic noise in magnitude.

Dror conducted studies on photon noise using Aiken's images and found that while photon noise leads to a slight increase in relative error, its contribution is small compared tov that of pattern noise suggesting that the performance of a velocity estimation system based on Reichardt correlators depends primarily on responses to pattern noise [Dror, 1998].

The accuracy in the estimation of the motion parameters depends largely on the field of view [Koenderink and van Doorn, 1987]. When the diameter of the receptive field is small, the nature of the field line pattern becomes unclear and there can be no decision of whether a field is due to translation or due to rotation or a mixture of both [Koenderink and van Doorn, 1987]. The tangential neurons on the fly brain are sensitive to the typical optic flow patterns generated during self-motion. It was found that a simplified linear model of these neurons which performs sampling from large receptive fields can be used to estimate self-motion from optic flow [Franz and Chahl, 2002].

In this chapter, different spatial sampling techniques are experimented in a correlation based system to understand their effect on pattern noise. It was found that a circular array of EMDs reduces pattern noise compared to a rectangular grid of EMDs or randomly selected EMDs for measuring rotational velocity, indicating that the circular method of sampling the detectors may be optimal for creating yaw detectors. In this study, the effects of saturation and adaptation are not in order to simplify the model.

6.2 Presence of Pattern Noise

Differences in the structure of scenes result in variations in the correlator response termed as the pattern noise. In order to clearly demonstrate this, physiological and modelling experiments were done at a high speed and at a low speed by repeating the same stimulus at the same and different initial positions.

6.2.1 Electrophysiology methods

Males of the fly *Eristalis tenax* were captured in the wild and the membrane potential of wide-field motion detecting neurons (HS cells) was recorded intracellularly using standard techniques [O'Carroll et al., 1997].

Stimuli were presented on a CRT display (Flatron 915 FT Plus, LG Electronics) at a frame rate of 200 Hz at 640 by 480 pixel resolution and mean luminance of 41 cd/m^2. A conventional OpenGL graphics card (geForce 3 Ti 200, nVidia) was used with Vision Egg stimulus generation software by author Andrew Straw (www.visionegg.org). The stimuli used were the same digital panoramas used for the modeling experiments. These panoramas were presented on the inside of a virtual cylinder, which was centered on the fly and perspective distorted according to the fly's calibrated 3D position and orientation relative to the screen. Motion simulated pure yaw rotations by spinning the virtual cylinder about the dorsal-ventral axis of the fly's head. The display subtended approximately 90° horizontally from midline to the lateral portion of the animal, thus stimulating most of the receptive field of HS neurons [Strausfeld and Seyan, 1985; Hausen, 1982c].

6.2.2 Electrophysiology results

By aligning responses in time across several initial stimulus positions, both pattern noise and intrinsic noise is largely removed from the mean, but both contribute to the standard deviation ('time aligned' in Figure 6.1 and Figure 6.2). When responses are aligned by stimulus position ('phase aligned'), the mean is subject to pattern noise while the standard deviation indicates only intrinsic noise. A comparison of the mean between time aligned and phase aligned responses shows the presence of pattern noise on real HS cells in our experimental setup. The larger standard deviation of the time aligned response compared to the phase aligned response indicates that pattern noise greatly exceeds intrinsic noise in HS cells at 850°/s . Unfortunately, our recording durations were too short to perform

the same phase alignment at 100°/s, but inspection of the individual responses indicates that pattern noise is less significant at this slower speed.

Figure 6.1 and Figure 6.2 shows the response of an HS neuron to the same image, with the image presented moving at 850 degrees per second and 100 degrees per second respectively for 3 seconds at eight different initial 'phases', each 45 degrees apart. The mean and standard deviation of the time varying response are given to illustrate intrinsic noise. The part (a) of both the images shows the response at phase zero and part (b) of Figure 6.1 and Figure 6.2 shows the response in each of these 8 configurations. Then in part (c), we average the resulting 8 curves in two ways. The first is 'time aligned' where the response is aligned by stimulus time. Because the initial phase is different in each trial, phase dependent noise (pattern noise) disappears in the average, just as if we had simulated a 360 degree EMD array. The second is 'phase-aligned', where we align each phase-delayed image with the data obtained for phase zero, so that it is the image position that is averaged, rather than time. By comparing the standard deviation of the time aligned responses to that of the phase aligned responses, we can see that 'pattern noise' dominates intrinsic noise. To eliminate noisy high frequency signal components, signals were low pass filtered before averaging. From the two figures, it is seen that the pattern noise is more prominent at higher velocities than lower velocities.

6.2.3 Modelling results

The above electrophysiological experiment is repeated using our model in an attempt to model pattern noise. Our elaborated EMD model [Rajesh et al., 2005b] is designed with an EMD averaging window, the same angular size as the computer screen used for the physiological experiment, and the simulation is run at the same speed as done in the physiological experiment (850°/s and 100°/s) using the same eight phases.

Figure 6.3 and Figure 6.4 shows the the modelling results at 850°/s and 100°/s respectively. The results at phase zero and at each different initial phase are shown and it can be seen that it is similar to the physiological results clearly showing the pattern noise present. Since intrinsic noise is totally absent in the model, the noise present is mainly caused by the pattern noise. Then as carried out before, the resulting 8 curves are averaged in two ways, (i) time aligned and (ii) phase aligned. In the time aligned method, the normal average of the each of the eight response removes the pattern noise as the response is aligned by stimulus time. In the phase aligned method, each phase delayed response is aligned by shifting each response along the x-axis with the data obtained for phase zero

and then we average it. Now the pattern noise is still present because in this way, it is the image position that is averaged rather than time, and we can still see the noise as the noise is locked in the stimulus position.The response of the model agrees well with the physiological results clearly demonstrating that the pattern noise is indeed a major source of temporal response variation in real HS recordings. It is also seen from both physiological and modelling results that pattern noise is present more significantly at higher speeds.

6.3 Receptive Field Shape and Pattern Noise

Our hypothesis is that the large strip-like nature of the HS cell receptive field, which is fully circular if summation at the steering muscles is taken into account, plays a functional role in encoding velocity. Because the circular array of EMDs is parallel to the preferred direction of motion, a purely circular optical flow pattern (i.e. yaw for an HS neuron) would produce the same temporal sequence of stimulation of individual photoreceptors. Hence if a high contrast feature is present within the receptive field at one point in time, it will also be present at future times. By summing across such an array, pattern noise is largely reduced.

In order to test this hypothesis, simulations were conducted with a panoramic yaw stimulus with a fixed velocity, for 2 seconds. The test was carried out for 4 different panoramic images (these images are shown in Figure 6.5) and 12 different speeds for each. The matrix of EMD outputs is then sampled in one of 4 ways.

The four ways in which the output of the EMDs are sampled

1. the output of all the EMDs are used which is called as total, as all the EMDs are used here,

2. the output of a randomly selected rectangular array of 15×16 EMDs are taken in this case,

3. the output of randomly selected 240 EMDs from the total sample of EMDs are used,

4. And finally the response of a randomly selected circular row of 240 EMDs is taken.

Each case is clearly explained in the subsections below.

6.3.1 Total Sample Of All EMDs

Figure 6.6 shows the total array of EMDs (240 × 37) from which the output of motion detection is logged. The inter-ommatidial angle of 1.5 degrees, which separates each EMD is not shown in this figure. Each box here represents an EMD. The total array is the sum (across the space dimension) of all EMDs. Here the outputs of the all EMDs (240 × 37 EMDs) is sampled. In this method, the outputs of all the EMDs are averaged over time where as in the forthcoming sections only selected EMDs are used from the total array and the effect of each of these sampling techniques on the pattern noise is noted by calculating the relative error which is clearly explained in the next section.

6.3.2 Square Lattice

A randomly selected rectangular grid of 15 × 16 EMDs (i.e. 240 total) is used here. Figure 6.7 shows a sample of randomly selected rectangular array of EMDs from which the output of the motion detection is noted. The inter-ommatidial angle of 1.5 degrees, which separates each EMD is not shown in this figure. Each box here represents an EMD and the shaded box represents the EMDs selected from the total array. Since sampling of a square array of detectors gives a similar response as that of a rectangular array, a rectangular array is also called a square lattice as is in this paper.

6.3.3 Randomly Sampled Array

In this method, the sample consisted of 240 EMDs selected randomly from the 240 × 37 array. Figure 6.8 shows a sample of randomly selected EMDs from the total array of 240 × 37 EMDs. The inter-ommatidial angle of 1.5 degrees, which separates each EMD is not shown in this figure. Each box here represents an EMD and the shaded box represents the EMDs selected for correlation.

6.3.4 Circular Sampled Array

Finally a circular linear array with just 240 EMDs, representing one of the 37 possible rows of EMDs, selected at random, is used. Figure 6.9 shows a sample of a randomly selected single row of 240 EMDs from which the output of the motion detection is noted. The inter-ommatidial angle of 1.5 degrees, which separates each EMD is not shown in

this figure. Each box here represents an EMD and the shaded box represents the EMDs selected for correlation.

6.4 Sampling Results

Figure 6.10 shows the error bars and responses obtained from the simulation of a rectangular, random, row and total array of EMDs. Part (a) of this figure shows the responses and relative error obtained from the simulation of a rectangular array (16×15) array of 240 EMDs. As can be seen the errors in this case indicate that there is more pattern noise when the sampling is done with a rectangular array of EMDs. Part (b) gives the error bars that are obtained from the simulation of 240 EMDs selected randomly from the total array of 240×37 EMDs. Because of the random selection of the EMDs, some error can be seen with this kind of sampling. Part (c) shows the error bars that are obtained from the simulation of a circular array of 240 EMDs. It is seen that the error in this case is reduced and is similar to using the whole total array of 240×37 EMDs indicating effective encoding of velocity can be done using even a small number of a circular detectors and finally part (d) shows the error bars obtained from the simulation of total array of 240×37 EMDs. In all four of the above curves, the correlator outputs are given in arbitrary units.

Figure 6.11 shows the mean correlator simulated response for the square, circular, random and total array of EMDs for image 1 moving at a constant velocity of 500 degrees per second. It is seen that the sampling of circular detector array gives a stable output similar to the total sample of EMDs whereas it can be seen that the results of the square or rectangular lattice is the worst with the presence of pattern noise seen more in this kind of sampling than with the random or circular array.

The error measure used here called the relative error defined by Dror as, $E_{\text{rel}} = E_{\text{abs}}/\overline{R}$, where the absolute error (E_{abs}) is the difference between the actual response and the expected response [Dror, 1998]. The expected response is the mean response value that is given by $\overline{R}$. For a given set of images, moving at a given velocity, the mean response $\overline{R}$ is calculated by averaging the response of the wide field correlator at all points in the selected sampled space and sampled time. The relative error for the same set of responses is found by dividing their standard deviation by the mean.

The mean relative error is then calculated for each case (square, circular, random and total) by finding the average relative error for about 20 iterations on all the four images at 12 different speeds and is plotted as shown in Figure 6.12. It is seen that the relative

error of the circular sample of EMDs is less than that of the random sample and the rectangular detector arrays and is much closer to the relative error of the total EMD array. This result suggests that motion detection and velocity estimation at yaw velocity can be done effectively using just a circular array of 240 EMDs.

The error bars in the Figure 6.10 and the mean correlator response for each case, as shown in Figure 6.11 also indicate that the circular sample of EMDs reduces errors and pattern noise more than the randomly sampled array and rectangular or square grid of EMDs [Rajesh et al., 2005c].

6.5 Role of Natural Image Statistics

Natural images are used as stimuli for the experiments. Natural images are far from random and show a large degree of structure. This structure can be described by the statistics of the image source, and can be considered as prior knowledge. Therefore a certain amount of image data is predictable and thus redundant [Ruderman, 1994; van der Schaaf and van Hateren, 1996]. The visual system appears to be optimized to take advantage of the statistical properties of natural images using specific optimization criteria of redundancy minimization, maximisation of information transmission, sparseness of the neural coding and minimising reconstruction error, demonstrating that simple optimisation principles combined with knowledge of image statistics can predict visual processing strategies which are found in nature [Ruderman, 1997]. Recent development in statistical modelling, along with powerful computational tools, have enabled researchers to study more sophisticated statistical models for visual images and to use them to test the efficient coding hypothesis for both individual neurons and populations of neurons [Simoncelli and Olshausen, 2001; Clatworthy et al., 2003; Petrov and Zhaoping, 2003]. The results of various studies done on the statistics of natural images all draw the same conclusion that the power spectrum of natural images tends to depend on the spatial frequency f in the form of $\frac{1}{f^2}$ [Tolhurst et al., 1992; Ruderman, 1994]. Even when individual images deviate from this power law, the differences seem to be confined to large scales [Langer, 2000]. Similar findings extend to the time domain [Dong and Attick, 1995; van Hateren, 1997].

Measured power spectra are one of the most prominent sources of natural image statistics, which are obtained by computing the Fourier transform of an image and multiplying each element of the transform by its complex conjugate. Averaging over horizontal and vertical orientations gives power as a function of frequency. The power spectra of natural images

are generally similar. The $\frac{1}{f^2}$ property of natural images is generally not obtained for random images (such as for example, unbiased random noise images, which produce a flat power spectrum). The $\frac{1}{f^2}$ spectral slope of natural images implies that natural images are statistically scale-free [Abbott, 2001]. That is, images viewed through different angular scales (through lenses of different focal lengths) have similar statistics [Field, 1987; Dror et al., 2000; van der Schaaf and van Hateren, 1996; Ruderman, 1994; Ruderman and Bialek, 1994; Ruderman, 1997].

Studies carried out by Dror reveal that similarities between natural image power spectra lead to predictable peak response velocities and to similarities in the shapes of the velocity response curves for different natural images [Dror et al., 2001]. The primary difference between the curves, their overall amplitude, results from contrast differences between images. In order to use mean correlator response as a reliable indicator of velocity, the visual system needs to compensate for these contrast variations [Dror et al., 2000; Dror, 1998]. One possibility is contrast saturation early in the motion detection pathway, which eliminates significant differences in contrast [Dror, 1998] and alternatively some form of adaptation (contrast adaptation) in the visual system, may work to remove contrast difference between the images [Harris et al., 1999; Rajesh et al., 2002].

6.6 Conclusion

In this paper, we have conducted a study on pattern noise using physiological recordings together with modelling results and we performed quantitative analysis of the effects of receptive field shape on pattern noise, measured as 'relative error'. This error is defined as the residual variance over time in the response due to the local structure of the image, divided by the mean (DC level) of that response. The assumption here is that the DC level of the summed outputs of an EMD (elementary motion detector) array represents the desired velocity signal, while time-variation due to the pattern structure is noise, so that the relative error is effectively a measure of signal-to-noise ratio in the context of the task of velocity coding. It is found that the circular receptive field shape reduces pattern noise compared to an array of rectangular or randomly selected EMDs.

It can also be seen that the circular linear array achieves a very similar variance in velocity estimation to that of the *full* EMD set, and (depending on speed) 10 or more times lower than the other two strategies for sampling EMDs. It is hence concluded that a circular, linear detector array may be the optimal way to sample outputs from EMDs of the

correlator type, at least with respect to reduction of pattern noise when estimating yaw velocity. Taken together with the underlying physiological experiments that are being carried out in our lab, this suggests that a small detector array with as few as 240 pixels could provide robust estimates of speed under natural image conditions. Further, it would be interesting to pursue similar tests on more complex optic flow.

In the next chapter, the affect of the presence of compressive non-linearity on pattern noise is shown.

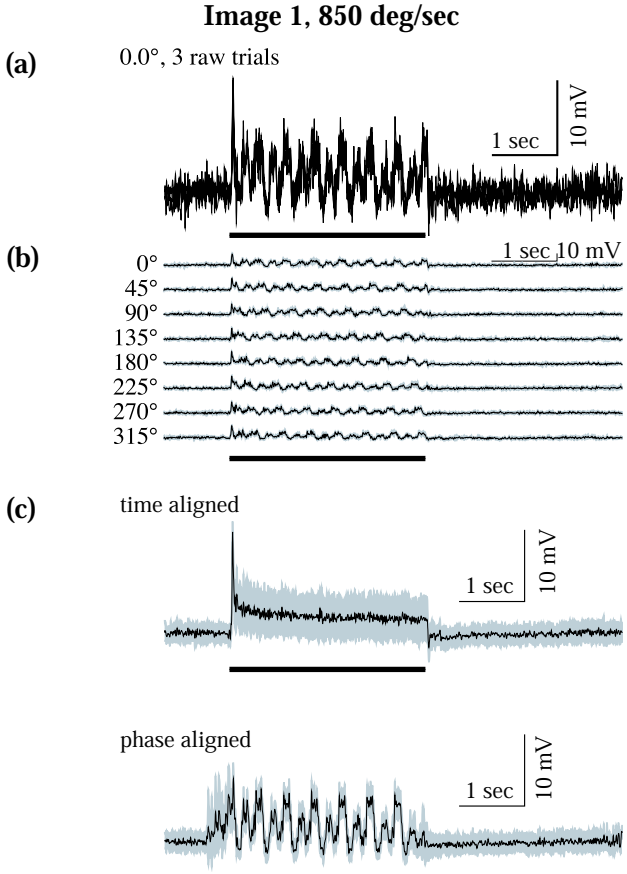

Figure 6.1. Presence of pattern noise in the response of a fly neuron with the image presented moving at high velocity. This figure shows the effect of 'pattern noise' on response of a single fly wide-field motion detecting neuron (HS cell) to image 1 (see Figure 6.5), with the image presented moving at $850°/s$ at 8 different initial phases, each 45 degrees apart. For each combination of velocity and image, graph (a) shows three responses to identical experimental conditions. Graph (b) shows the mean of at least three responses at each of several stimulus positions. Graph (c) show the mean response in black and the standard deviation of the response around the mean in gray. The 'time aligned' response shows the individual responses at each initial position averaged without compensating for position change. For the $850°/s$ motion, responses were also 'phase aligned' by compensating for the position change before averaging. To eliminate noisy high frequency signal components, signals were low pass filtered before averaging.

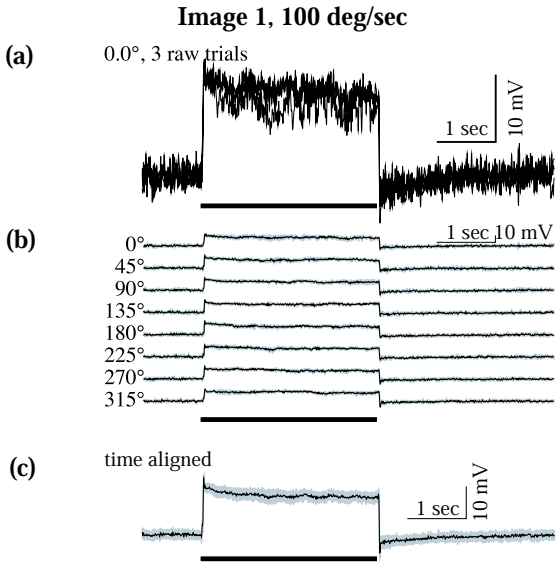

Figure 6.2. Presence of pattern noise in the response of a fly neuron with the image presented moving at low velocity. This figure shows the effect of 'pattern noise' on response of a single fly wide-field motion detecting neuron (HS cell) to image 1 (see Figure 6.5), with the image presented moving at $100°/s$ at 8 different initial phases, each 45 degrees apart. For each combination of velocity and image, graph (a) shows three responses to identical experimental conditions. Graph (b) shows the mean of at least three responses at each of several stimulus positions. Graph (c) show the mean response in black and the standard deviation of the response around the mean in gray. The 'time aligned' response shows the individual responses at each initial position averaged without compensating for position change. Unfortunately, our recording durations were too short to perform the same phase alignment at $100°/s$, but inspection of the individual responses indicates that pattern noise is less significant at this slower speed. To eliminate noisy high frequency signal components, signals were low pass filtered before averaging.

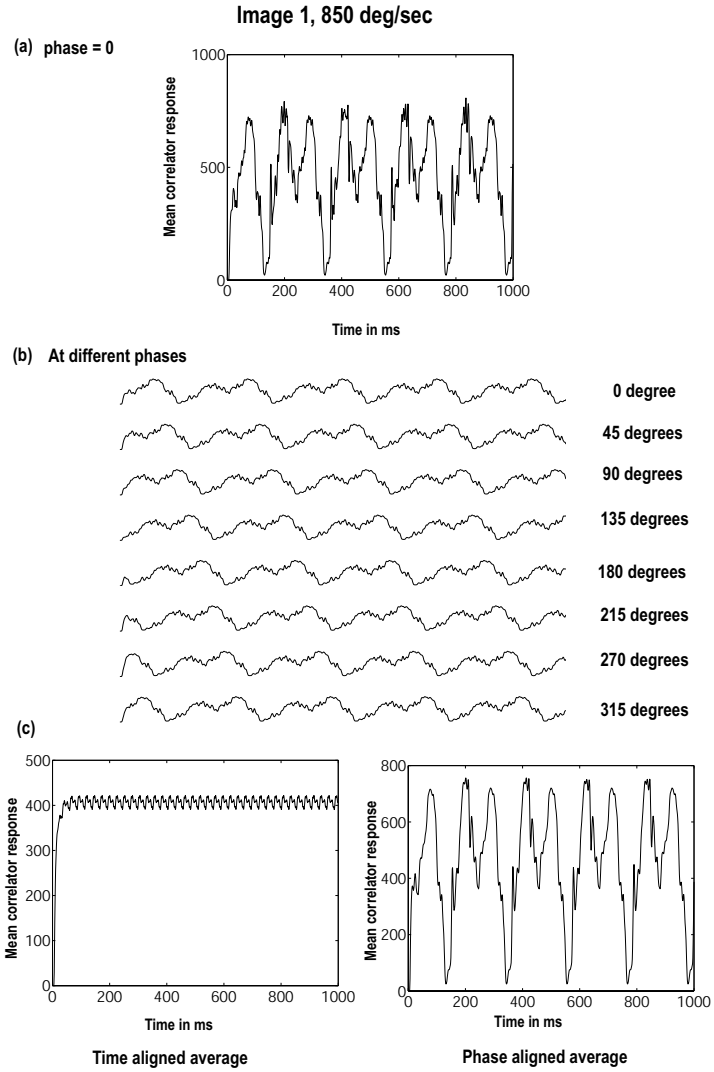

Figure 6.3. Presence of pattern noise in the simulated response of our elaborated model with the image presented moving at high velocity. This figure shows the simulation results obtained by running the model at $850°/s$ using the same image, at 8 different initial phases, each 45 degrees apart, as done in the physiological experiment shown in Figure 6.1. Graph (a) of the graph shows the response at phase zero. Graph (b) shows the response to each of the 8 configurations. Graph (c) shows the response averaged in the time aligned way. In the time aligned method, the normal average of the each of the eight response removes the pattern noise as the response is aligned by stimulus time. The results obtained agree with the physiological data, demonstrating that pattern noise acts as a major source of temporal response variation in HS recordings.

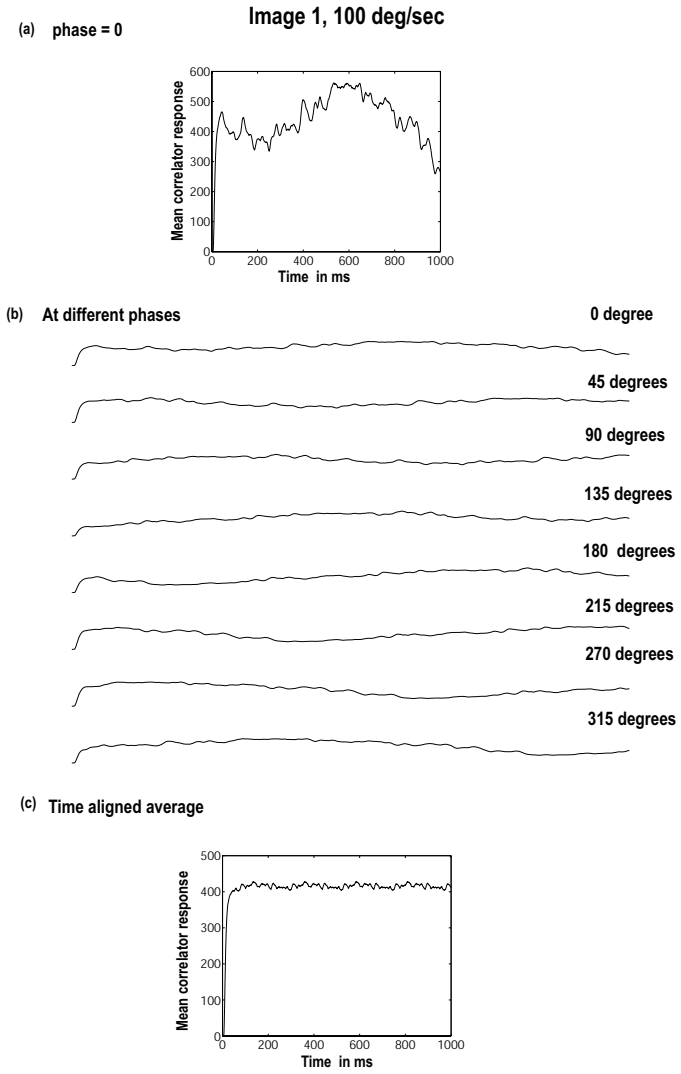

Figure 6.4. Presence of pattern noise in the simulated response of our elaborated model with the image presented moving at low velocity. This figure shows the simulation results obtained by running the model at $100°/s$ using the same image, at 8 different initial phases, each 45 degrees apart, as done in the physiological experiment shown in Figure 6.2. Graph (a) of the graph shows the response at phase zero. Graph (b) shows the response to each of the 8 configurations. Graph (c) shows the response averaged in two ways, time aligned and phase aligned. In the time aligned method, the normal average of the each of the eight response removes the pattern noise as the response is aligned by stimulus time. In the phase aligned method, each phase delayed response is aligned by shifting each response along the x-axis with the data obtained for phase zero and then we average it. Now the pattern noise is still present because in this way, it is the image position that is averaged rather than time, and we can still see the noise as the noise is locked in the stimulus position. The results obtained agree with the physiological data, demonstrating that pattern noise acts as a major source of temporal response variation in HS recordings.

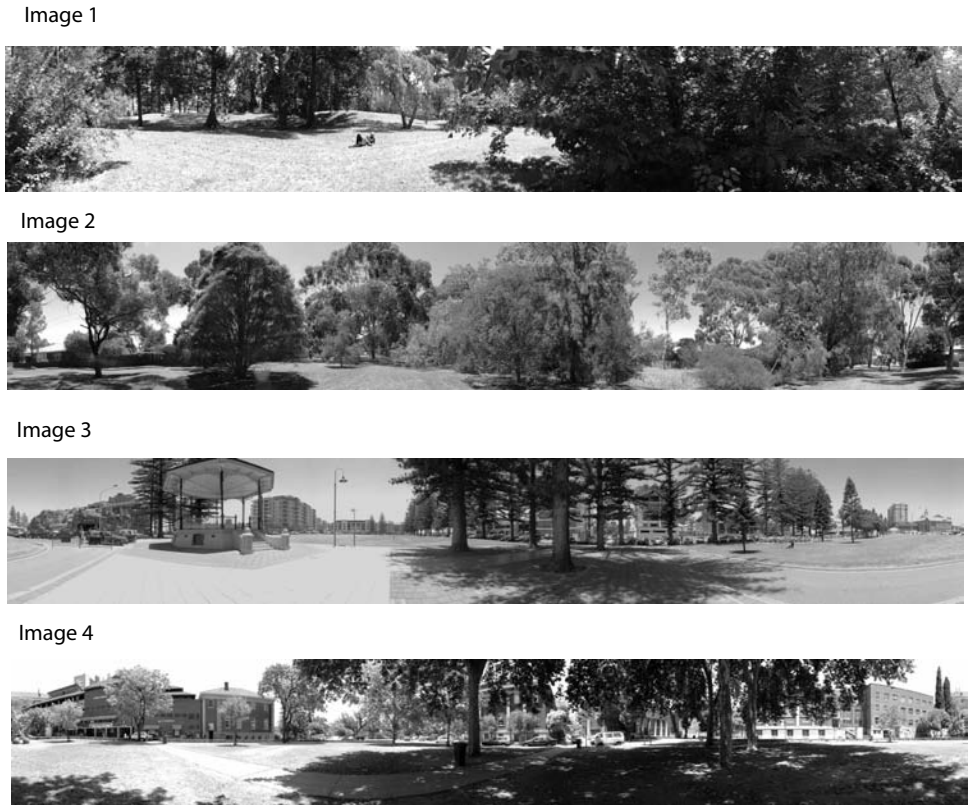

Figure 6.5. The panoramic images used as stimuli to our model. A panorama of the image is formed by 'warping' 12 image tiles at 30° intervals to remove lens distortions and then by wrapping its ends together using Apple Quicktime VR software on a Macintosh computer.

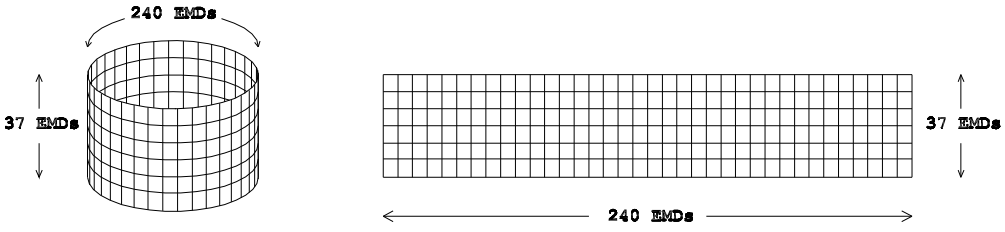

Figure 6.6. Total EMD array. This diagram shows the total array of EMDs (240×37) from which the output of motion detection is logged. The inter-ommatidial angle of 1.5 degrees, which separates each EMD is not shown in this figure. Each box here represents an EMD.

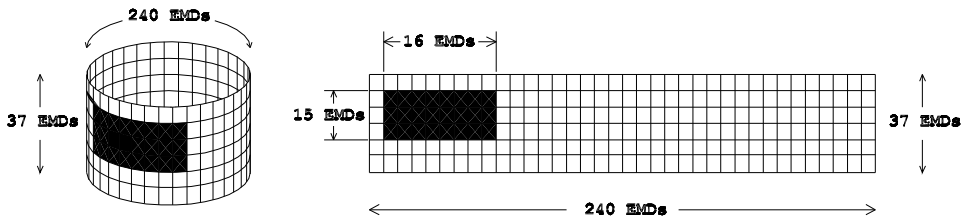

Figure 6.7. Rectangular EMD sample. This diagram shows a sample of randomly selected rectangular array of EMDs from which the output of the motion detection is noted. The inter-ommatidial angle of 1.5 degrees, which separates each EMD is not shown in this figure. Each box here represents an EMD and the shaded box represents the EMDs selected from the total array.

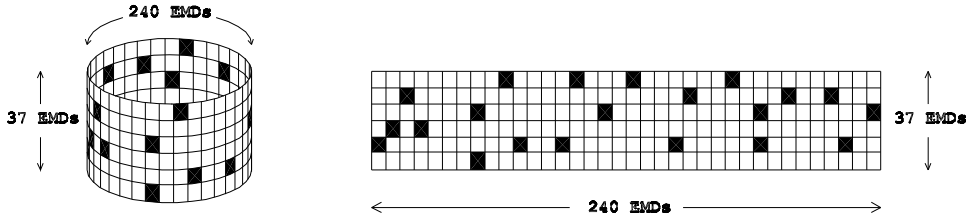

Figure 6.8. Random EMD sample. This diagram shows a sample of randomly selected EMDs from the total array of 240×37 EMDs. The inter-ommatidial angle of 1.5 degrees, which separates each EMD is not shown in this figure. Each box here represents an EMD and the shaded box represents the EMDs selected for correlation.

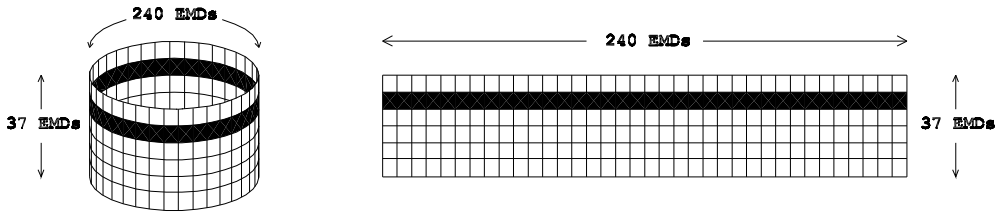

Figure 6.9. Row or circular EMD sample. This diagram shows a sample of a randomly selected row of 240 EMDs from which the output of the motion detection is noted. The inter-ommatidial angle of 1.5 degrees, which separates each EMD is not shown in this figure. Each box here represents an EMD and the shaded box represents the EMDs selected for correlation.

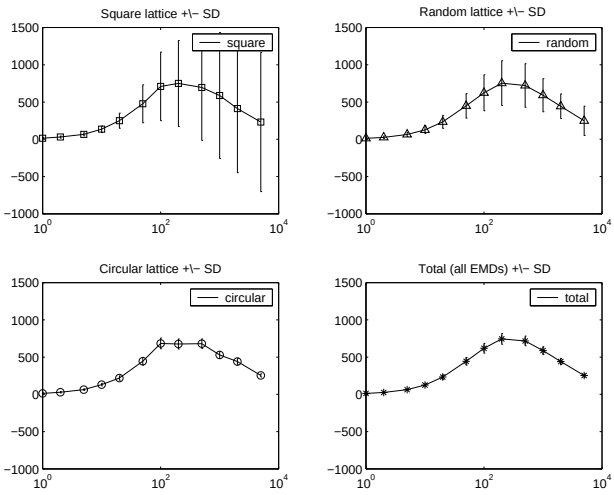

Figure 6.10. Error bars and responses obtained from the simulation of a rectangular, random, row and total array of EMDs. (a) Responses and relative error obtained from the simulation of a rectangular array (16×15) array of 240 EMDs. As can be seen the errors in this case indicate that there is more pattern noise when the sampling is done with a rectangular array of EMDs. (b) The error bars are obtained from the simulation of 240 EMDs selected randomly from the total array of 240×37 EMDs. Because of the random selection of the EMDs, some error can be seen with this kind of sampling. (c) The error bars are obtained from the simulation of a circular array of 240 EMDs. It is seen that the error in this case is reduced and is similar to using the whole total array of 240×37 EMDs indicating effective encoding of velocity can be done using even a small number of a circular detectors. (d) The error bars obtained from the simulation of total array of 240×37 EMDs. In all four of the above curves, the correlator outputs are given in arbitrary units.

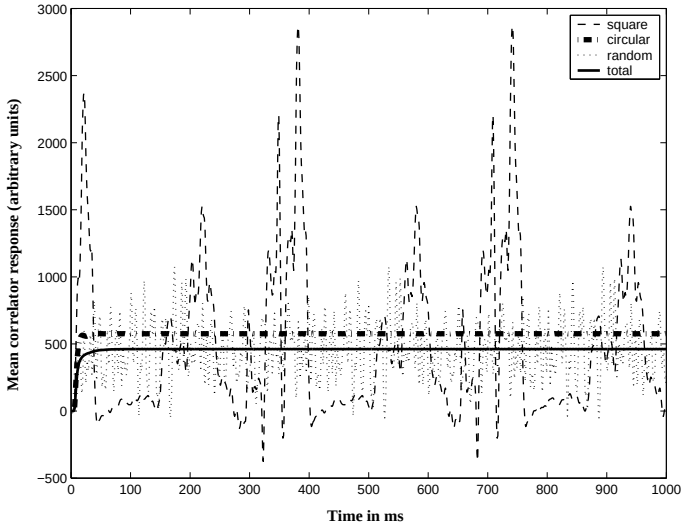

Figure 6.11. Simulated mean correlator response for square, circular, random and total array of EMDs. The mean correlator simulated response for the square, circular, random and total array of EMDs for image 1 moving at a constant velocity of 500 degrees per second is shown here. It is seen that the sampling of circular detector array gives a stable output similar to the total sample of EMDs.

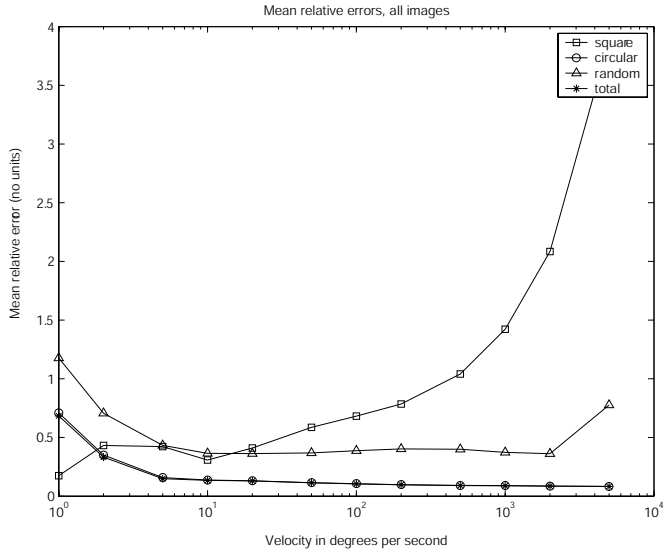

Figure 6.12. The relative errors obtained from the simulation of the square, random, circular and total array of the EMDs. The relative error defined by Dror as, $E_{\text{rel}} = E_{\text{abs}}/\overline{R}$, where the absolute error (E_{abs}) is the difference between the actual response and the expected response [Dror, 1998]. The mean relative error is calculated by averaging the relative error obtained for about 20 iterations on all the four images at 12 different speeds. It is clearly seen that the mean relative error for the circular linear array of EMDs is much lower and is closer to that of the total array of detectors.