Do gradient models of primate vision perceive coherent motion in dynamic Glass patterns?

Ed Long, CoMPLEx

Supervisors: Prof. Alan Johnston & Prof. Peter McOwan

Word count: 4920

January 26, 2007

Abstract

Dynamic Glass patterns are random image sequences whose local structure gives rise to an apparent global motion. They are of interest in the field of computational visual processing: previous papers predict that they should contain no low-level motion signals, conjecturing that the percept is due to a mechanism in the ventral pathway. In this essay, I summarise a number of motion perception models and describe processing a dynamic Glass pattern stimulus using a gradient model of motion perception. I found that, in the sequences I used, there was evidence for a coherent low-level motion for particular parameters but my analysis opposed the view that this was the sole factor inducing an apparent motion.

Contents

1 Introduction

1.1 The visual system

Humans perceive their surroundings using information from light which passes through the pupil of each eye and creates a two-dimensional image on the retina—the layer of light-sensitive cells at the back of the eye. A pinhole camera makes a convenient analogy but the manner in which the brain extracts information from the retinal image is very different from the development of a photo. Firstly, the optic nerve does not have sufficient bandwidth to send a signal from every retinal cell so an aggregation of data must take place. Secondly, we have the ability to perceive motion and form an estimate of a three dimensional image, so the brain must have a means to interpret a two dimensional image in order to create a three-dimensional map. A key part of this is the combination of the information from the left and right eye.

Physiological and psychophysical experiments have enabled us to roughly map the visual pathway. Direct recording of electrical impulses in neurons has been employed primarily in macaque monkeys to map activity in reaction to various stimuli. For example, displaying a particular colour may cause an individual neuron to fire more rapidly; in another neuron the same effect may be observed when a drifting intensity grating matches a particular angle. In lesion studies, areas of interest are damaged and it is noted which visual abilities are lost.

The above two techniques cannot ethically be used to study human perception but the development of fMRI has allowed noninvasive study of brain activity. Also, brain damage in stroke and head-trauma patients gives useful, though less comprehensive, information about the location of visual processing areas.

Crudely (see [2] Chapter 3), light hitting the retina is coded into an electrical signal which passes down the optic nerve via the *optic chiasm* (where the signals from the right and left eye cross over into the opposite half of the brain) to the *lateral geniculate nucleus* (LGN). The LGN then relays the signal to the *visual cortex* situated at the back of the head, where the majority of the processing takes place.

Terminology varies in naming areas of the cortex. The first stage of processing takes part in the *striate cortex*, a name derived from its striped appearance, frequently referred to as *V1*. From V1 the visual pathway branches into a number of other regions. These are collectively referred to as the *extrastriate cortex*.

Branches into and between areas of the extrastriate cortex are diverse but experiments suggest they roughly form two functional pathways. These are called the *ventral pathway* and the *dorsal pathway*. The dorsal pathway appears to be more important in motion perception whereas the ventral pathway builds a detailed picture of the surroundings and responds to geometrically complicated visual cues such as facial recognition.

Evidence that different features of a spatiotemporal image are processed by different areas is seen in patients with brain damage where, for example, they have the ability to navigate their surroundings but cannot recognise faces or are blind on one side but still have the ability to react to movement in their blind spot ("blindsight"). A number of further examples are cited in [2].

1.2 Computational models

The understanding of human vision is inextricably tied to the development of computing both as a method of generating precise and quantifiable stimuli and in terms of implementing algorithms designed to model the visual pathway.

David Marr in his 1982 book [9] first put forward a framework for computational visual processing based on a modular set of algorithms which extract, in stages, a three-dimensional representation of the surroundings from the two-dimensional image on the retina.

In order to implement an algorithm, the information available must be coded mathematically. In Landy & Movshon's collection of articles [8], Adelson & Bergen describe a generalised function termed the *plenoptic function*, $P(\theta, \phi, \lambda, t, V_x, V_y, V_z)$ or $P(x, y, \lambda, t, V_x, V_y, V_z)$, giving all possible information about the surroundings viewed in any direction (θ, ϕ) and from any point (V_x, V_y, V_z) over all time, t. The wavelength of the light is λ and the output of the function is an individual intensity value for each of the seven coordinates. Given that light focuses on a point from all directions, polar coordinates are more appropriate but if we restrict the angles available to the observer then the coordinates can be mapped to a Cartesian grid (x, y), which is easier to implement in computer models.

The power of Marr's approach lies in stressing computational function rather than underlying physiology in order to determine possible algorithmic methods to extract features from visual input and then to test hypotheses generated by the models against physiological and psychophysical data.

1.2.1 Correlation models

An intuitive idea for motion detection is that we notice a moving object at two distinct points in time and calculate its velocity $\mathbf{v}$ (bold face denotes a vector) by dividing the change in position by the time period. ie.

$$\mathbf{v} = \frac{\Delta \mathbf{x}}{\Delta t} \tag{1}$$

Whilst this is an attractive theory, we must still define how an observer chooses which elements of the visual field to track and which elements in a subsequent time frame to correlate them to. The number of points in an image that can be reliably tracked is limited so, in the case of noisy or particularly complex visual input, it would be a very difficult task to perform this correlation at each time step.

1.2.2 Linear filters and motion detectors

An alternative approach is to find a way to compute local velocities for all points in the image, rather than tracking particular features. This involves treating a moving image as a continuous spatiotemporal image, rather than a series of discrete time frames although, in implementation, discrete jumps between pixels or frames can be smoothed by applying a blurring function.

Models of this type have biological motivation: the receptive fields of neurons are two-dimensional with inhibitory and excitatory regions so we can define filters with similar properties to these to extract information from the moving image.

As a simplification of the plenoptic function mentioned above, consider the intensity map I in (x, y, t)-space. We can then apply filters to compare any two of the variables and seek to extract the orientation of the intensity map. This is useful for detecting the shape of objects but, more usefully, it reduces the the problem of detecting direction of motion to extracting the orientation of the intensity map in (x, t) or (y, t) space. *Motion as orientation* is a paradigm frequently referred to in the literature.

A classic motion detector example is the model proposed by Reichardt [10], in which the (x, t) filter contains two excitatory regions at a particular orientation. With a bank of these, each at a slightly different orientation, the filter with the strongest output tells us the local orientation of the spatiotemporal image and hence the local velocity.

A second approach, developed by Adelson [1], uses nonoriented filters based on *Gabor functions*—products of sine or cosine functions and Gaussian functions—as a primary input. Combination of inputs of this type can be used to extract a leftward and rightward *motion energy*. The difference of these gives the *opponent energy*, which is then normalised to give a measure which estimates local velocity[1].

[1] As pointed out in [2], [1] and a number of other sources, choices of particular linear filters as the basis of the above two models can transform a motion energy model into a Reichardt motion detector model or vice versa. For this reason, the two model types are often treated as having equivalent features.

1.3 Gradient models

The class of model considered in this essay is the *gradient model*. This uses partial derivatives of various orders of the intensity map to calculate a measure of velocity.

It may not be intuitive that a neural mechanism for calculating partial derivatives is biologically plausible, but Koenderink [4] describes how convolution of the spatiotemporal image with partial derivatives of Gaussian functions can be used to give blurred spatial and temporal partial derivatives of the intensity to any degree. These derivatives of Gaussians also have qualities very similar to neural receptive fields.

The base Gaussian blurring function is of the form:

$$G(x,t;\sigma) = \frac{e^{-\frac{x^2+t^2}{4\sigma}}}{\sqrt{4\pi\sigma}} \tag{2}$$

where σ determines the spread of the function.

We denote partial derivatives by $G_x = \frac{\partial G}{\partial x}$, $G_{xx} = \frac{\partial^2 G}{\partial x^2}$, $G_t = \frac{\partial G}{\partial t}$ etc. Then, denoting the *convolution*,

$$\int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau, \tag{3}$$

of f and g by $f \otimes g$ we have:

$$\frac{\partial^n}{\partial x^n}[I(x,t) \otimes G(x,t;\sigma)] = \frac{\partial^n I}{\partial x^n} \otimes G = I \otimes G_{x\ldots x} \tag{4}$$

$$\frac{\partial^n}{\partial t^n}[I(x,t) \otimes G(x,t;\sigma)] = \frac{\partial^n I}{\partial t^n} \otimes G = I \otimes G_{t\ldots t} \tag{5}$$

where $G_{x\ldots x}$ denotes $\frac{\partial^n G}{\partial x^n}$ and likewise for t.

Specifically, then, we have $\frac{\partial I}{\partial x} \otimes G = I \otimes G_x$ and $\frac{\partial I}{\partial t} \otimes G = I \otimes G_t$, so convolving the intensity map with a derivative of the blurring function, G, returns a blurred spatial or temporal derivative of the intensity map.

We can use this to calculate local velocities: if our feature is moving at velocity v in the x direction, then, as detailed in [2], chapter 8, the intensity is a function $I(x-vt)$. Then, since $I_x = I'(x-vt)$ and $I_t = -v \cdot I'(x-vt)$ we have that

$$v = -\frac{I_t}{I_x} \tag{6}$$

Or, in terms of the linear filters:

$$v = \frac{I \otimes G_t}{I \otimes G_x} \tag{7}$$

Despite its neatness, this approach has weaknesses in that a zero denominator implies an infinite velocity and, when the denominator is small, the calculation is much more sensitive to a noisy signal.

It is unlikely that both the first and second derivatives of the intensity map will be zero. As outlined in [6], given the relation $I_t = -v \cdot I_x$ we can differentiate with respect to x again to get $I_{xt} = -v \cdot I_{xx}$. In general:

$$\frac{\partial}{\partial t}\frac{\partial^n}{\partial x^n} = -v \cdot \frac{\partial^{n+1}}{\partial x^{n+1}} \tag{8}$$

Coding our partial derivatives in vectors $\mathbf{x} = (x_1, x_2) = (I_x, I_{xx})$ and $\mathbf{t} = (t_1, t_2) = (I_t, I_{xt})$ we wish to find a scalar u such that $\mathbf{t} = -u\mathbf{x}$. For the best fit, we must minimise the distance $\|\mathbf{t} + u\mathbf{x}\|$.

Taking the square, we have

$$A = \|\mathbf{t} + u\mathbf{x}\|^2 = \sum_i (t_i + ux_i)^2 \tag{9}$$

At a minimum, we must have

$$\begin{aligned} \frac{dA}{du} &= 2\sum_i x_i(t_i + ux_i) && (10)\\ &= 2\mathbf{x}\cdot(\mathbf{t} + u\mathbf{x}) && (11)\\ &= 2(\mathbf{x}\cdot\mathbf{t} + u\mathbf{x}\cdot\mathbf{x}) && (12)\\ &= 0 && (13) \end{aligned}$$

and hence

$$u = -\frac{\mathbf{x}\cdot\mathbf{t}}{\mathbf{x}\cdot\mathbf{x}} = \frac{I_x I_t + I_{xx} I_{xt}}{I_x^2 + I_{xx}^2} \quad (14)$$

Note that the vectors can include an arbitrary number of derivatives, with robustness increasing with each derivative taken.

Here we have only considered one spatial variable as described in Johnston *et al.*'s 1992 paper [6] but, in the 1999 paper [7], this approach is generalised to two spatial dimensions. In [5], Johnston & Clifford show that models of this type are able to account for a number of motion illusions (the *fluted square wave illusion, reverse phi illusion* and *Pantle illusion*), opposing the theory that there are multiple mechanisms for detecting different types of motion.

1.4 Glass pattern stimuli

Glass patterns (named after physiologist Leon Glass after he introduced them in his 1969 paper, [3]) are images consisting of randomly positioned pairs of dots. A randomised *seed* pattern is first generated, typically within a circular boundary, and then partner dots are superimposed on the image after transforming the original pattern by a particular geometric rule. This could be a rotation, translation, expansion or a combination of the above.

Despite their random pattern, the local cues give rise to a percept of global structure.

Dynamic Glass patterns are image sequences, each frame of which is a Glass pattern. Although there is no global motion in the sequence, viewers often report a strong global motion percept aligned with the orientation of the dot pairs, but with ambiguous direction. In fact, observers can "choose" the direction of apparent motion: in a rotational Glass pattern an observer can make the pattern appear to rotate either clockwise or anticlockwise.

Ross *et al.* [12] asked participants to view rotational Glass patterns with 100 dot pairs and rate the strength of the rotational percept as they varied the angle of dot separation and frame rate. They found that the optimal parameters were close to the optimal parameters for seeing real global motion. They also found that, for short dynamic Glass pattern sequences, observers were unable reliably to differentiate the apparent motion percept from sequences of similar patterns containing real global motion.

Ross [11] later studied which parameters affected the perceived speed and direction in various types of Glass pattern and Glass line pattern (in which the dot pairs are replaced by short line segments). With human observers he found that an increase in the length of the lines led to an increased apparent speed of motion but increased dot separation in a dot-pair Glass pattern did not. He argued that the local form of the Glass pattern line or dipole is interpreted as a motion streak by the visual system: that is, the local form of the stimulus is what gives the apparent motion percept.

Ross also investigated combining Glass pattern stimuli with coherent motion signals. He found that the Glass pattern orientation influenced the perceived direction of motion but that this influence, or *motion deflection*, was greatly reduced when the angle of coherent motion differed from the dot-pair orientation by more than 30 degrees. He also noted that, whereas in a superposition of two coherent motion signals with different orientations observers can detect two separate velocities—an effect called *motion transparency*—no motion transparency is observed in a superposition of a coherent signal and a Glass pattern.

The perception of Glass patterns has also been studied at the cortical level. Smith *et al.* [13] measured the responses of neurons in the V1 area of macaque monkeys when they were made to observe a translational Glass pattern image sequence. They determined the optimal spatial period of each cell by tuning a drifting sine grating of variable wavelength and found that the the Glass pattern produced a spatial tuning curve similar to that predicted by their model. The cells displayed clear selectivity for Glass pattern orientation but the responses were much smaller than for a drifting grating.

I chose to investigate whether there were any detectable low-level motion signals in the case of rotational Glass pattern sequences, using Johnston *et al.*'s multi-channel gradient model [7]. I also considered how the strength of the percept could be quantified and how it was affected by the number of dot pairs in the image and the angle of dot pair separation. If a strong low-level motion signal is evident in patterns of this type, it would add weight to the evidence in [5] for a single motion-processing mechanism.

2 Method

I generated a dynamic Glass pattern stimulus using the code included in Appendix A.1. The stimulus consisted of 64 frames of 128 by 128 pixel bitmaps with white dot pairs inside a circle of radius 64 pixels on a black background. The parameters that I varied were the number of dot pairs in each image and the angle used in the rotation to generate the pattern of partner dots (the angle of separation). The script generated random coordinates in x and y, discarding pairs that were outside the radius of the circle. It then transformed the coordinate vectors to obtain the coordinates of the partner dots and then used these to change values at those positions in a 128 by 128 zero matrix to 1. The point (64,64) was marked with a grey dot so that the centre of rotation could be checked in each frame. The original script generated a movie so I could check the rotational percept and choose parameters that gave rise to a strong impression of rotation.

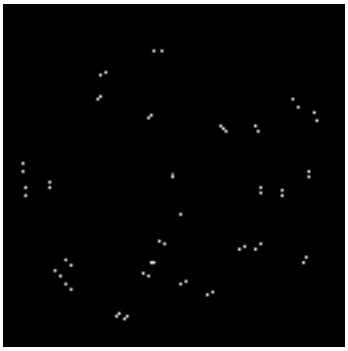
To detect low-level motion signals in the stimulus, I used a computer implementation of Johnston's multi-channel gradient model (MCGM). For the purpose of generating input for the MCGM, instead of an animation I generated a series of numerically indexed individual bitmaps.

I chose to generate image sequences with combinations of 3, 6 and 9 degrees of dot separation and with 30, 60 and 90 dot pairs (in total nine image sequences). See Figure 1 for examples of single frames from a number of the image sequences.

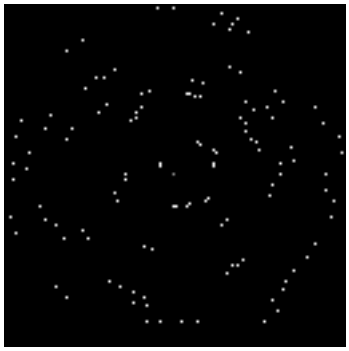
3 Results

I ran the MCGM process using the nine dynamic Glass pattern image sequences as an input. The process generates binary files with a calculated angle and speed for each pixel. It also displays these as bitmaps with angle or speed coded in an intensity map. Secondly, it applies a speed threshold mask to the raw data to generate a masked image which displays only areas moving above a certain speed. This shows the moving components of the image more clearly and, in the case of the angle data, codes angle of motion using a hue colour map so the direction of motion can be matched to a colour wheel displayed in a border around the image. Examples of these are shown in Figure 2 in the case with 30 dot pairs and 3 degrees angle of separation. As the model samples over time by comparing stacks of 22 images, one image sequence generates 41 of each type of output file.

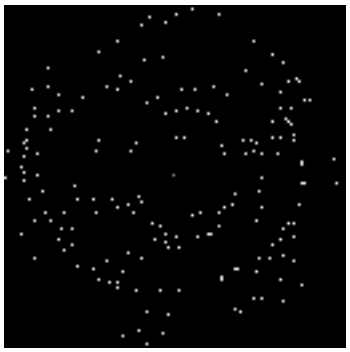
Figure 2(a) shows the apparent direction of motion of all regions moving above the threshold speed. For a strong rotational motion signal, there should be a bias of red and cyan dots lying on the line through the centre from left to right and a bias of green and blue dots lying on the line through the centre from top to bottom. This does not appear to be the case at first glance.



(a) 30 dot pairs with 3 degrees separation

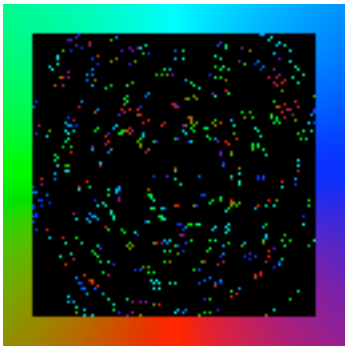


(b) 60 dot pairs with 6 degrees separation

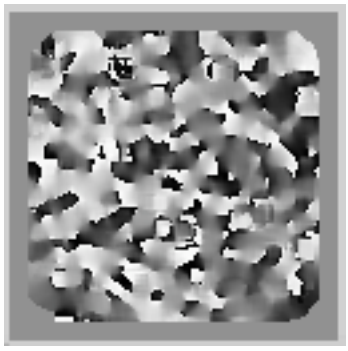


(c) 90 dot pairs with 9 degrees separation

Figure 1: Some examples of single Glass pattern frames

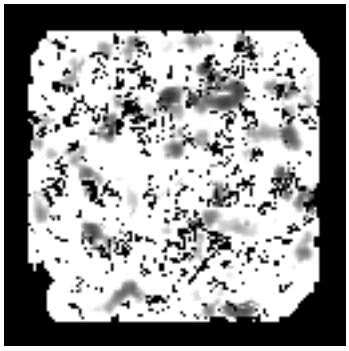


(a) Bitmap showing angle of movement (masked)



(b) Bitmap showing angle of movement (raw data)

(c) Bitmap showing local velocities (masked)



(d) Bitmap showing local velocities (raw data)

Figure 2: Examples of MCGM output images

4 Analysis

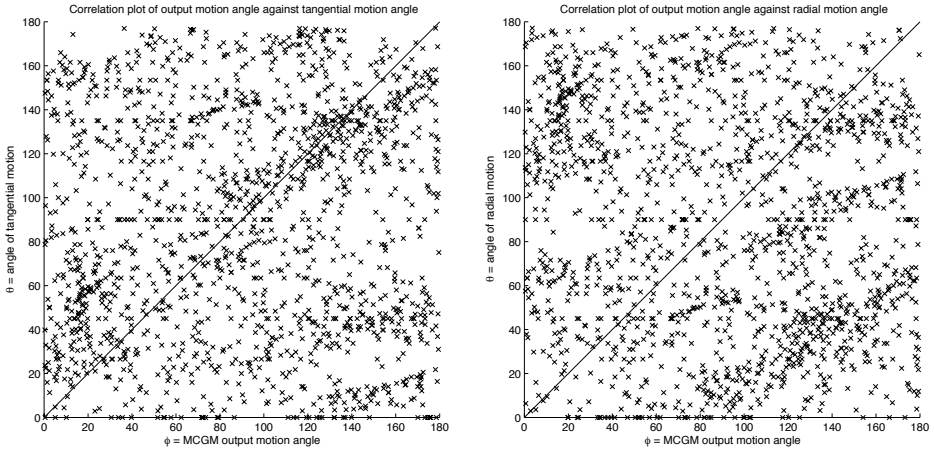
Using the script in Appendix A.2 I extracted the angle of motion, ϕ, for each pixel in the first binary output file (based on the first 22 images in the sequence) and converted the angles into degrees, measured anticlockwise from the horizontal around the centre of the image.

In order to quantify the strength of a rotation percept, I compared the apparent strength of a tangential motion (rotation) and a radial motion (expansion/contraction). To do this, I constructed a second matrix with each entry being the azimuth, θ, of the corresponding pixel. For a purely expansional motion, the direction of motion is along the line of the azimuth so θ and ϕ are equal at all points. For a purely rotational motion, ϕ is equal to $\theta \pm 90^\circ$, depending on the direction of rotation.

Since the rotational percept can be perceived as either clockwise or anticlockwise, I considered all angles *modulo* 180°. This reduces the analysis simply to whether motion is a tangential or radial movement.

The script in Appendix A.3 plots the angle data against the expected angle of a purely radial or purely tangential motion. Since sampling all points in the image produced an unnecessarily dense plot and since the matrix included a zero border which skewed the results, I took the angles from points within a 40 by 40 square at the centre of the image.

An example of the scatter plots generated for the sequence with 30 dot pairs and 3 degrees separation is shown in Figure 3. A full table of plots for all image sequences is given in Appendix B.



(a) Tangential motion

(b) Radial motion

Figure 3: Scatter plots for image sequence with 30 dot pairs and separation of 3 degrees

By inspection it is clear that there is no strong correlation with either a tangential or radial motion. In order to see whether there was at least a slightly stronger indication of tangential motion I used the script in Appendix A.4 to calculate the Pearson correlation coefficient for each of the scatter plots. The results are shown in Table 1.

At first glance we can see that the correlation between ϕ and a tangential motion is at least mostly positive, though for 60 dot-pairs separated by 9 degrees there is zero correlation and there is negative correlation for two of the other results. The correlation is not strong: at most almost 8% in the image sequence with 30 dot-pairs separated by 6 degrees. The percept seems roughly to be strongest with a smaller number of dot-pairs and a small angle of separation.

The correlation between ϕ and a radial motion is, conversely, negative in seven of nine cases. Surprisingly, the strongest correlation of all observations is the negative radial correlation in the case with

Image sequence	Tangential motion	Radial motion
30 dot pairs, 3 degrees	0.0644	-0.1140
30 dot pairs, 6 degrees	0.0758	-0.0026
30 dot pairs, 9 degrees	0.0201	-0.0813
60 dot pairs, 3 degrees	0.0410	-0.0286
60 dot pairs, 6 degrees	-0.1037	0.0704
60 dot pairs, 9 degrees	0.0000	-0.0178
90 dot pairs, 3 degrees	0.0326	-0.0399
90 dot pairs, 6 degrees	0.0083	0.0124
90 dot pairs, 9 degrees	-0.0311	-0.0838

Table 1: Pearson correlation coefficient for tangential vs. radial motion for all image sequences

30 dot-pairs separated by 3 degrees at just over 11%. A rotational motion, however, should be weakly negatively correlated to the radial motion, so this supports the existence of a very weak low-level motion signal.

In order to quantify the significance of the results, I also calculated the p-values for each of the observations: that is, the probability that randomly generated data points would produce the same correlation value as observed in Table 1. The p-values are shown in Table 2 as well as an indication of whether the correlation was positive or negative.

Image sequence	Tangential motion	Radial motion
30 dot pairs, 3 degrees	*0.0100* (+)	*0.0001* (-)
30 dot pairs, 6 degrees	*0.0024* (+)	0.9172 (-)
30 dot pairs, 9 degrees	0.4217 (+)	*0.0011* (-)
60 dot pairs, 3 degrees	0.1011 (+)	0.2529 (-)
60 dot pairs, 6 degrees	*0.0001* (-)	*0.0048* (+)
60 dot pairs, 9 degrees	1.0000 (+)	0.4768 (-)
90 dot pairs, 3 degrees	0.1925 (+)	0.1106 (-)
90 dot pairs, 6 degrees	0.7401 (+)	0.6202 (+)
90 dot pairs, 9 degrees	0.2137 (-)	*0.0008* (-)

Table 2: Level of significance (p-values) for the correlation coefficients in Table 1

A statistically significant correlation is typically interpreted as being one with a p-value lower than 0.05: i.e. there is less than a 5% probability that random data will produce a correlation of that strength or greater. These entries are shown in italic in Table 2. We can see that, although all of the correlation coefficients were smaller than 12%, the large number of data points (1600) means that a number of the results are statistically significant.

For each of the image sequences with 30 dot pairs, at least one of the correlation coefficients with tangential or radial motion is significant under these criteria. Furthermore, all of the significant data in this subset support the existence of a tangential motion.

For the 60 dot-pair image sequence, the only significant data support radial rather than tangential motion.

For the 90 dot-pair sequence only one correlation coefficient is significant: it supports a tangential motion but the correlation with tangential motion is negative so it is unclear what conclusions we can draw from this result.

5 Discussion

Given the strength of the motion percept that can be observed in dynamic Glass patterns it is surprising that, at first sight, the correlation between the measured local velocities against a model rotational pattern is so small. Statistical measurements, however, show that there is a statistically significant bias towards a low-level rotational motion signal in image sequences with particular parameters. With the dot-separation angles used in the image sequences, a sparser arrangement of dots separated by relatively smaller angles appears to give a stronger impression of rotational motion.

An interesting consideration is whether the rotational percept depends on an association made between the two dots in the pair or a *false matching* between a dot in one frame and one in a subsequent frame. If the local geometry of the Glass pattern image sequence had some bias in which the closest dot in a subsequent frame were more likely to lie at a particular orientation to any given dot in the first frame, then this should give a bias towards a low-level motion signal in that direction.

First, consider the case with a random arrangement of N unpaired dots in each time frame. All lie within a circle of radius R. We then choose one of these dots and ask what the probability is of a dot in the next time frame lying within a circular neighbourhood of radius $r < R$ of the first dot.

The probability that each dot will lie within this neighbourhood is the quotient of the areas of the greater and smaller circle: $\frac{\pi r^2}{\pi R^2} = \left(\frac{r}{R}\right)^2$.

Hence the probability that at least one dot lies in this circle is given by:

$$\begin{aligned} P(\textit{at least one dot}) &= 1 - P(\textit{no dots}) && (15) \\ &= 1 - \left(1 - \left(\frac{r}{R}\right)^2\right)^N && (16) \end{aligned}$$

In the case of a Glass pattern image, we must also take into account the local geometric properties of the pattern. A dot lying within the neighbourhood could either be a seed dot or a partner dot. Hence the probability for any individual dot lying inside the neighbourhood is proportional to the area of the union of the original neighbourhood, C, and its pre-image D under the transformation: any dot could either be a seed lying in the neighbourhood or a partner dot of a seed lying in the pre-image.

If we assume that $r << R$ then any rotation or expansion will locally approximate a translation. Hence we calculate the union of the neighbourhood with its pre-image as the overlapping area of two circles of equal radius with centres offset by a small amount s in the positive x direction (Figure 4).

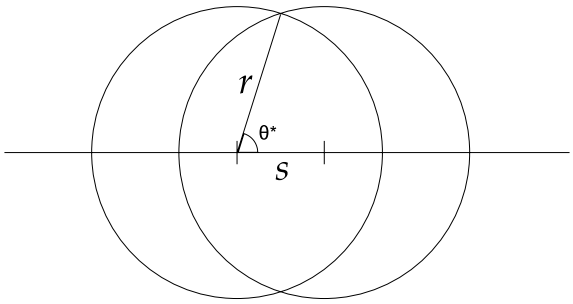


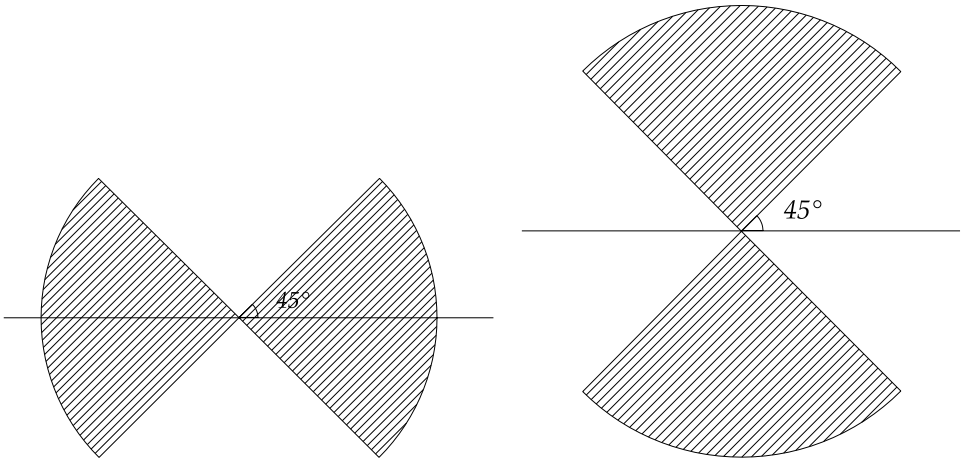
Figure 4: The union $C \cup D$

Then, if θ^* is the angle between the positive x-axis and the line passing through the centre of the left-hand circle and the point of intersection of the two circles, the area of the figure is:

$$A_{C \cup D} = 2\left(\pi r^2 - \theta^* r^2 + rs \sin\theta^*\right) \quad (17)$$

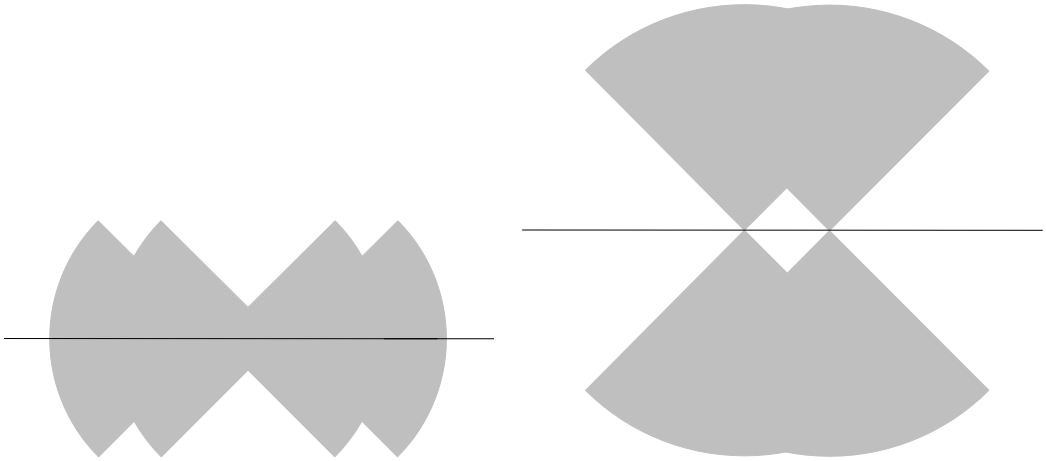
Note that $\theta^* = \arccos(\frac{s}{2r})$.

To investigate a vertical versus horizontal bias, we instead consider sectors of the neighbourhood with their pre-images. Dividing the neighbourhood into diagonal quadrants on a 45 degree angle gives the figures shown in Figure 5. The union of these with their pre-images is shown in Figure 6.



(a) Horizontal (b) Vertical

Figure 5: Quadrants containing horizontally and vertically oriented points



(a) Horizontal (b) Vertical

Figure 6: The union of the quadrants in Figure 5 and their pre-images.

We calculate the area A_H of the horizontally-oriented union of quadrants by decomposing it as shown in Figure 7. The area of the larger sector is $\frac{\pi r^2}{8}$. Application of the sine rule gives the value of the angle marked ϕ as:

$$\phi = \frac{\pi}{4} - \arcsin\left(\frac{s}{r}\sin\frac{\pi}{4}\right) \tag{18}$$

$$= \frac{\pi}{4} - \arcsin\left(\frac{s}{\sqrt{2}}\right) \tag{19}$$

Hence the area of the triangle is $\frac{1}{2}rs\sin\phi = \frac{rs}{2}\sin\left(\frac{\pi}{4} - \arcsin\left(\frac{s}{\sqrt{2}}\right)\right)$.

The area of the smaller sector is $\frac{(\pi-4\phi)r^2}{8}$. This gives the total area of the decomposition in Figure 7 as:

$$\frac{\pi r^2}{8} + \frac{rs}{2}\sin\left(\frac{\pi}{4} - \arcsin\left(\frac{s}{\sqrt{2}}\right)\right) + \frac{1}{2}\arcsin\left(\frac{s}{\sqrt{2}}\right) \tag{20}$$

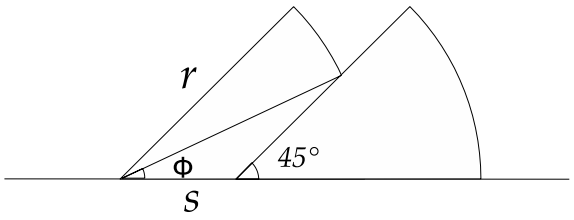


Figure 7: Decomposition of the horizontally-oriented quadrants.

Finally, superimposing four of the figures in Figure 7 and taking away the area of overlap: $\frac{s^2}{2}$ gives the total area:

$$A_H = \pi r^2 2 + 2rs\sin\left(\frac{\pi}{4} - \arcsin\left(\frac{s}{\sqrt{2}}\right)\right) + 2\arcsin\left(\frac{s}{\sqrt{2}}\right) - \frac{s^2}{2} \tag{21}$$

Calculating the area A_V of the vertically-oriented union is simpler. We take the total area of two overlapping circles (Equation 17) and subtract the area of two 90 degree sectors of the circle (on on each side) and the same quantity $\frac{s^2}{2}$ for the square at the centre. In all:

$$A_V = 2\left(\pi r^2 - r^2\arccos(\frac{s}{2r}) + rs\sin[\arccos(\frac{s}{2r})]\right) - \frac{\pi r^2}{2} - \frac{s^2}{2} \tag{22}$$

Using Mathematica, I calculated values for the ratio $\frac{A_H}{A_V}$ for $r = 5$ and s varying between 0 and 1 (Table 3). With $s = 0$, clearly the areas are equal but, as s increased, the ratio became smaller.

s	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
r	1.0000	0.9887	0.9773	0.9656	0.9537	0.9416	0.9292	0.9166	0.9037	0.8906	0.8771

Table 3

Since the probability of any individual dot lying in a region is proportional to the area of that region, these calculations show that increased dot separation leads to a decrease in bias towards a Glass pattern type motion. It implies that motion is more likely to be seen orthogonally to the orientation of the dot-pair than parallel to it. Although this supports the result that the increase in separation between the dots led to a weaker correlation with the expected motion, it contradicts the idea that a low-level motion signal arises from false matches between dots in subsequent time frames.

The evidence that we can draw from the above data to support a particular mechanism being solely responsible for a motion percept arising from dynamic Glass patterns is at best patchy. Due to time constraints I unfortunately only had time to run the model on these nine image sequences and only used a central sample from the first output file of each sequence to draw the scatter plots and calculate correlation coefficients. With more time, investigation of a wider range of the dot density and separation might have given a clearer picture of the directional bias. Another possible analysis technique would be to apply the velocity threshold used in creating the masked angle image before doing the same correlation, although there was no clear rotational bias visible in the masked angle bitmaps created by the MCGM (Figure 2).

The results, however, would seem to rule out the motion percept arising solely from low-level motion signals. The apparent motion was clearly visible watching the animations generated by the script for all tested values of the two parameters. This implies that another factor must influence this particular percept. The fact that the observer can choose the direction of apparent motion also shows that the percept cannot be entirely passive (i.e. data-driven) since the attention of the viewer is not a factor which is part of the stimulus.

In Bruce *et al.*'s description [2] of the visual pathway it stresses that the process is far from linear—that feedback loops feature at all levels of processing. Computational models of motion perception such as the MCGM are inherently data-driven: the percept of motion is derived from the bottom up with no addition of extraneous data. It is possible that the perception of motion could be biased by a feedback loop either from the information about local form, by an individual's memory, or both.

This type of feedback loop agrees with Ross's hypothesis [12] that local form can give rise to an apparent global motion: dot pairs or lines being interpreted as motion streaks. Another interesting consideration is whether the memory of form can bias perception of motion.

An observer's expectation can clearly affect our interpretation of an image. An example is the *reverse mask* illusion where a concave mask of a face appears to be convex since our expectation that faces are convex biases our construction of the three-dimensional map from the two-dimensional image. This mechanism could be important in motion perception as well as spatial perception. It is reasonable that experience might lead us to expect a pattern with a concentric circular form to rotate and one with a linear form to slide parallel to the direction of the lines. This adds another possible factor to the mechanism perceiving apparent motion of this type.

In conclusion, despite the small range of parameter values tested, the results from the model imply that dynamic Glass patterns do not contain low-level motion signals strong enough to justify the observed effect. Consideration of the probability of making false matches between timeframes opposes the hypothesis that motion signals could arise by this means. Whilst there is evidence that low-level motion signals exist for particular parameter values, the appearance of motion must be biased by input from areas of the brain other than those involved in motion detection.

A MATLAB code

A.1 Creating image sequence

```
n=90;               %Number of dot-pairs
theta=9*pi/180      %Angle of dot-pair separation

nframes=64;         %Number of frames in animation

for j=1:nframes
    dotimage=zeros(128);    %Create blank matrix
    dotimage(64,64)=0.5;    %Mark centre point
    v1=zeros(1,n);          %Initialise coordinate vectors
    v2=zeros(1,n);

    i=1;

    while(i<n+1)
        v1(i)=128*rand; %Choose random coordinates
        v2(i)=128*rand;
        if (64-v1(i))^2 + (v2(i)-64)^2 < 64^2 %Check that dot is inside circle
            i=i+1;
        end
    end

    xpos=v2-64;
    ypos=64-v1;
    xpos2=(xpos)*cos(theta) - (ypos)*sin(theta); %Rotate dots by angle theta
    ypos2=(ypos)*cos(theta) + (xpos)*sin(theta);
           %Wrap if rotated dot is outside image
    v1a=mod(64-ypos2,129);
    v2a=mod(xpos2+64,129);
```

```
    for i=1:n
        dotimage(ceil(v1(i)),ceil(v2(i)))=1;    %Add white dots at positions given by xpos, ypos
        dotimage(ceil(v1a(i)),ceil(v2a(i)))=1;  % "  "  xpos2, ypos2
    end

    if j<10
        cdigit=['00' num2str(j)];       %Assign suffix number for output files
    elseif j <100
        cdigit=['0' num2str(j)];
    else
        cdigit=num2str(j);
    end

    imwrite(dotimage,['dot' cdigit '.bmp']);    %Save numbered bitmap image

end
```

A.2 Reading binary output files

```
%Takes binary data and creates column vector with central sample of angles in degrees
f=fopen('AngOutput_128_X_128_float_0022.bin','r');
anglematrix=fread(f,[128,128],'float32');   %Reads binary data to 128 by 128 matrix
anglematrix=anglematrix(44:83,44:83);       %Samples from central 40 by 40 square in matrix
anglematrix=mod(anglematrix*180/pi,360);    %Convert radians to degrees: 0-360
phi=anglematrix(:);                         %Output matrix as vector
```

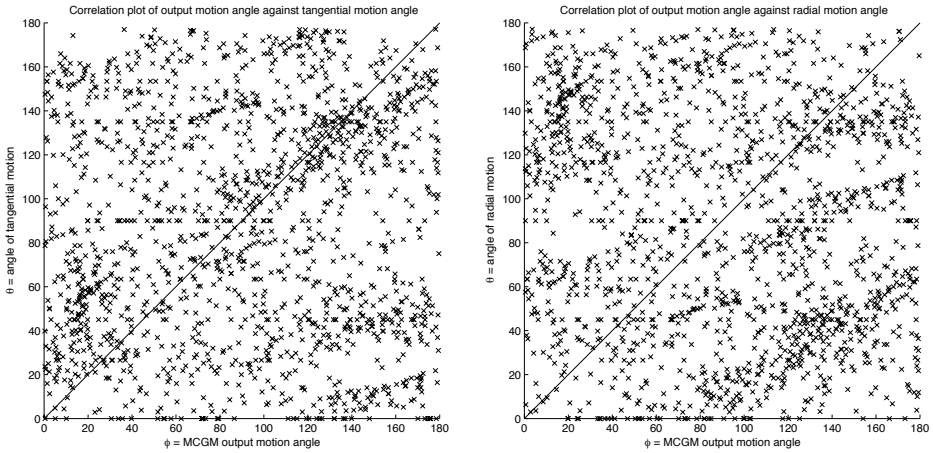
A.3 Creating scatter plots

```
%Sets up theta data and produces plots for tangential and radial motion
X=(1:40)'*ones(1,40);
thetaref=atan2(20-X,X'-20)*180/pi;
radial=mod(thetaref(:),180);
tangential=mod(thetaref(:)+90,180);
phimotion=mod(phi,180);

scatter(phimotion,radial,'x','k')       %Draws scatter plot of MCGM output against radial motion
xlabel('\phi = MCGM output motion angle')
ylabel('\theta = angle of radial motion')
title('Correlation plot of output motion angle against radial motion angle')
axis equal
axis([0 180,0 180])
hold on
plot(0:180,0:180,'-k')

figure                  %New figure for tangential motion

scatter(phimotion,tangential,'x','k')   %Draws scatter plot of MCGM output against tangential motion
xlabel('\phi = MCGM output motion angle')
ylabel('\theta = angle of tangential motion')
title('Correlation plot of output motion angle against tangential motion angle')
axis equal
axis([0 180,0 180])
hold on
plot(0:180,0:180,'-k')
```

A.4 Calculating correlation coefficients for tangential vs. radial motion

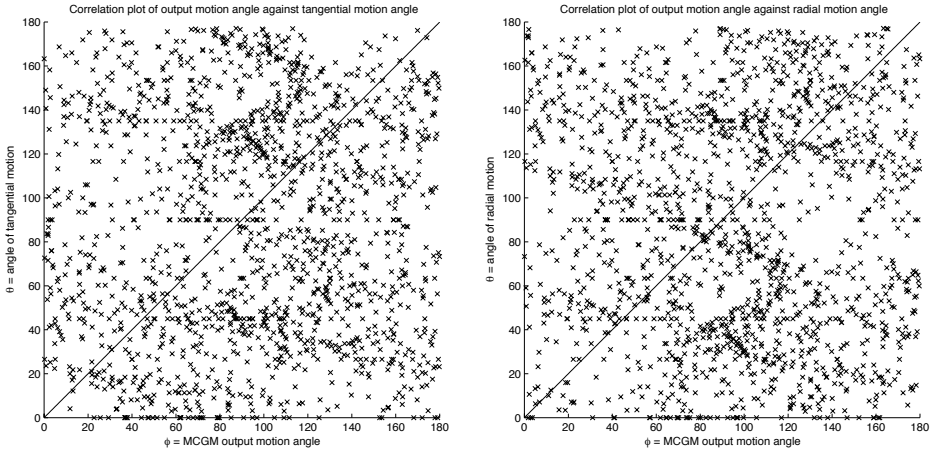
```
%Outputs Pearson correlation coefficient for tangential and radial motion
X=(1:40)'*ones(1,40);
thetaref=atan2(20-X,X'-20)*180/pi;  %Gives matrix with angle of pixel at each pixel location
radial=mod(thetaref(:),180);        %Modulo 180 as motion can be in either direction
tangential=mod(thetaref(:)+90,180); %Radial motion at 90 degrees to location
phimotion=mod(phi,180);
tangentialcor=corrcoef(phimotion,tangential)    %Calculates Pearson correlation coefficient
radialcor=corrcoef(phimotion,radial)
cortable(i,1)=tangentialcor(1,2);               %Enters coefficients into a table
cortable(i,2)=radialcor(1,2);
i=i+1                               %Increments counter for place in table
```

B Scatter plots



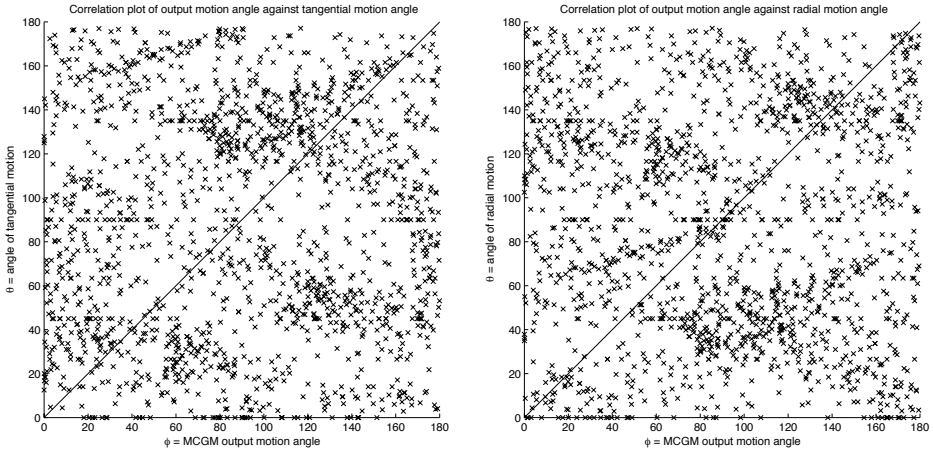
(a) Tangential motion (b) Radial motion

Figure 8: Scatter plots for image sequence with 30 dot pairs and separation of 3 degrees



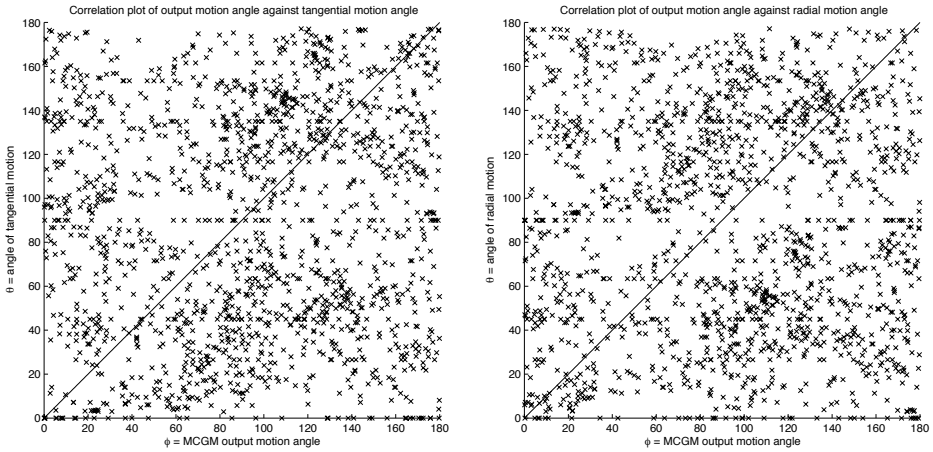
(a) Tangential motion (b) Radial motion

Figure 9: Scatter plots for image sequence with 30 dot pairs and separation of 6 degrees



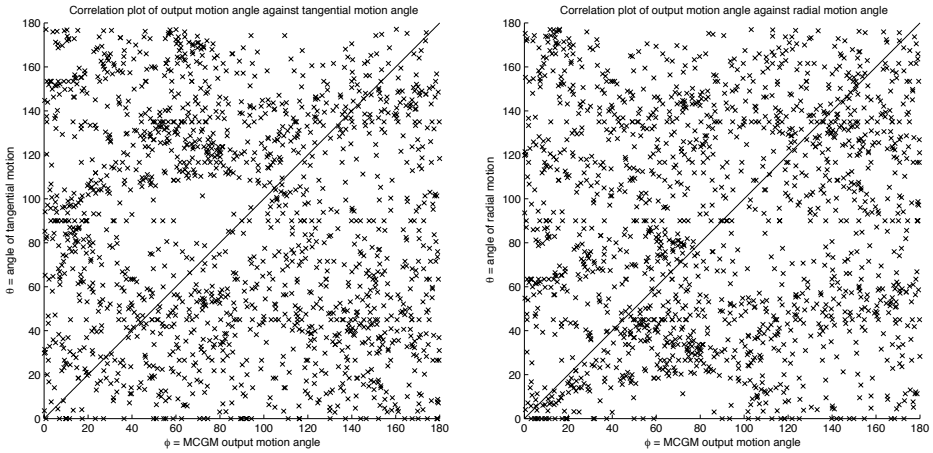
(a) Tangential motion (b) Radial motion

Figure 10: Scatter plots for image sequence with 30 dot pairs and separation of 9 degrees



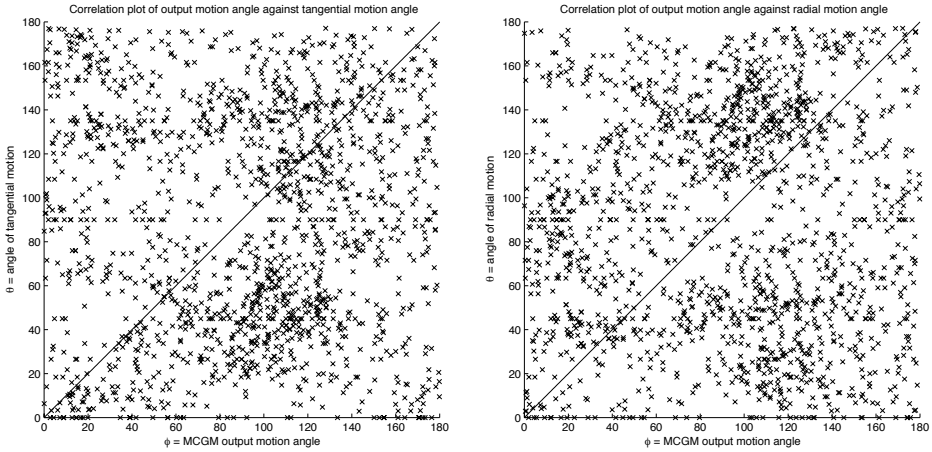
(a) Tangential motion (b) Radial motion

Figure 11: Scatter plots for image sequence with 60 dot pairs and separation of 3 degrees



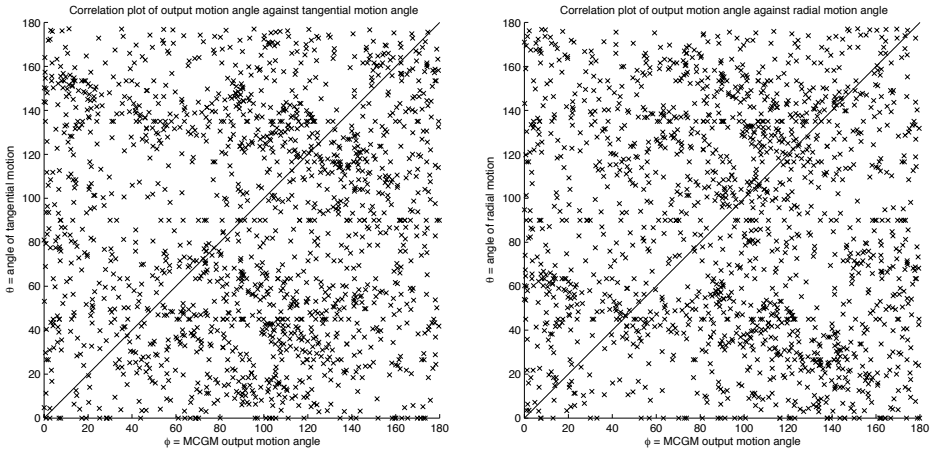
(a) Tangential motion (b) Radial motion

Figure 12: Scatter plots for image sequence with 60 dot pairs and separation of 6 degrees



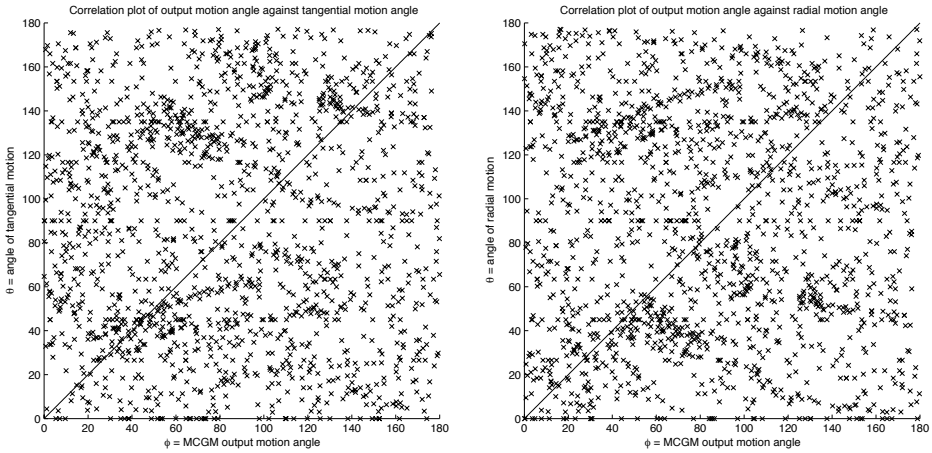
(a) Tangential motion (b) Radial motion

Figure 13: Scatter plots for image sequence with 60 dot pairs and separation of 9 degrees



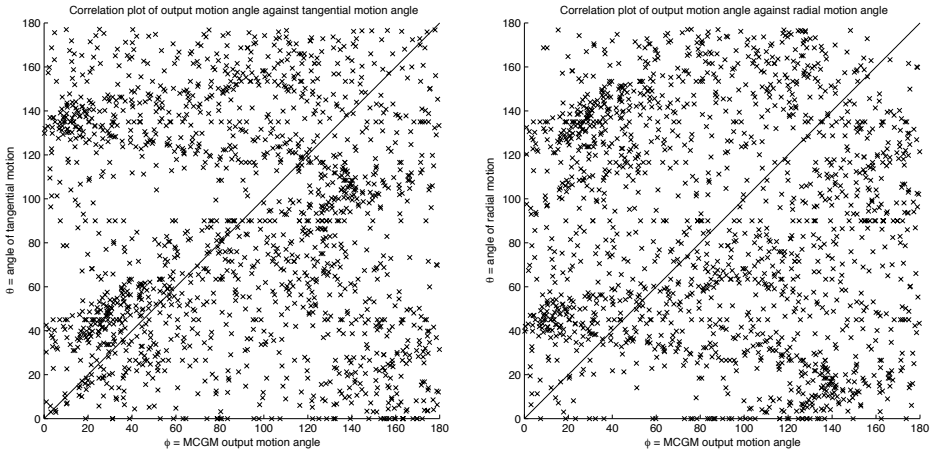
(a) Tangential motion (b) Radial motion

Figure 14: Scatter plots for image sequence with 90 dot pairs and separation of 3 degrees



(a) Tangential motion (b) Radial motion

Figure 15: Scatter plots for image sequence with 90 dot pairs and separation of 6 degrees



(a) Tangential motion (b) Radial motion

Figure 16: Scatter plots for image sequence with 90 dot pairs and separation of 9 degrees

References

[1] Edward H. Adelson and James R. Bergen. Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am.*, 2:284–299, 1985.

[2] Vicki Bruce, Patrick R. Green, and Mark A. Georgeson. *Visual Perception*. Psychology Press, 2003.

[3] Leon Glass. Moiré effect from random dots. *Nature*, 223:578–580, 1969.

[4] J.J.Koenderink and A.J. van Doorn. Representation of local geometry in the visual system. *Biological Cybernetics*, 55:367–375, 1987.

[5] A. Johnston and C.W.G. Cliffford. A unified account of three apparent motion illusions. *Vision Res.*, 35(8):1109–1123, 1995.

[6] A. Johnston, P.W. McOwan, and H.Buxton. A computational model of the analysis of some first-order and second-order motion patterns by simple and complex cells. *Proc. R. Soc. Lond. B*, 250:297–306, 1992.

[7] Alan Johnston, Peter W. McOwan, and Christopher P.Benton. Robust velocity computation from a biologically motivated model of motion perception. *Proc. R. Soc. Lond. B*, 266:509–518, 1999.

[8] Michael S. Landy and J. Anthony Movshon, editors. *Computational Models of Visual Processing*. MIT Press, 1991.

[9] David Marr. *Vision: A computational investigation into the human representation and processing of visual information*. W.H. Freeman & Co., 1982.

[10] W. Reichardt. Movement perception in insects. *Processing of optical data by organisms and by machines*, 1969.

[11] John Ross. The perceived direction and speed of global motion in Glass pattern sequences. *Vision Research*, 44:441–448, 2004.

[12] John Ross, David R. Badcock, and Anthony Hayes. Coherent global motion in the absence of coherent velocity signals. *Current Biology*, 10:679–682, 2000.

[13] Matthew A. Smith, Wyeth Bair, and J. Anthony Movshon. Signals in macaque striate cortical neurons that support the perception of Glass patterns. *Journal of neuroscience*, pages 8334–8345, 2002.