Complex Motion Processing
Richard Stewart Watson
University of London
Phd
2002
Table of Contents

Abstract 6
List of Figures and Tables 7

Chapter 1
Introduction

1.1 Abstract 26
1.2 Literature Review 27
1.2.1 The Utility of Motion Information 27
1.2.2 The Computational Approach 28
1.2.3 Modelling Optic Flow Analysis 32
1.2.4 Physiology 35
1.2.5 Psychophysics 37
1.2.5.1 Speed and Direction in Local Mechanisms 38
1.2.5.2 Speed and Direction in Global Mechanisms 40

Chapter 2
General Methods

2.1 Technical Issues 48
2.1.1 Sampling the Observer’s Psychometric Function 48
2.1.2 Stimuli 49
2.1.3 Data Processing 54
2.1.4 Model Fitting 57
2.1.5 Goodness-of-Fit 60
2.1.6 Confidence Limits on the Parameters Estimated 62
2.1.7 Software and Hardware 64
2.1.7.1 Software Code Generation 64
2.1.7.2 Hardware 64
Chapter 3
Direction Bandwidths of Rotation, Radial and Translation Motion Mechanisms

3.1 Abstract 72
3.2 Introduction 73
3.3 Experiment 1 77
3.3.1 Methods 80
3.3.2 Results 82
3.3.3 Discussion 86
3.4 Experiment 2 90
3.4.1 Methods 91
3.4.2 Results and Discussion 91
3.5 Experiment 3 100
3.5.1 Methods 104
3.5.2 Results 107
3.6 General Discussion 108

Chapter 4
The Dependency of Sensitivity on Spatial Frequency in Radial Flow

4.1 Abstract 112
4.2 Introduction 113
4.3 Experiment 1 117
4.3.1 Methods 117
4.3.2 Results 124
4.4 Experiment 2 126
Chapter 7
Do Relative Motion and Binocular Disparity Interact in the Detection of Shear Motion?

Chapter 8
Conclusions

Bibliography

Appendix 1
List of Abbreviations
ABSTRACT

When an observer moves, or objects in his field of view move, complex patterns comprising a heterogeneous range of speeds and directions are projected onto the retinas. Several classes of these optic flow patterns have been identified, including translation, rotation, radiation and shear. This thesis investigates the properties of motion-sensitive detectors in human vision using established psychophysical techniques.

The range of directions to which rotation and radial motion detectors are sensitive were measured. These ‘bandwidths’ were found to be narrower than estimated by previous investigators, a result that is incompatible with the notion that complex patterns are coded by a basis set of elementary components.

The size and spectral content selectivity of expansion detectors was measured. Surprisingly the visual system prefers pattern elements not to change size as they approach, contrary to real life. Similarly, optimal performance was found with constant spatial frequency, except at high speeds. We conclude that the spatial frequency tuning of low-level motion detectors places a fundamental limit on the detection of radial patterns.

A cue conflict paradigm was used to explore the exaggerated speed experienced with radial motion. Additional depth cues were ineffective in compromising the illusion, suggesting that it may not be due to a motion-in-depth interpretation, or that cue combination in the visual system weights relative motion more highly than other depth cues.

The smallest difference in speed that can be discriminated in rotation, radial motion and translation patterns was measured using two techniques. We conclude that thresholds are in keeping with existing literature, though results were inconsistent between techniques.

The meaning of some forms of motion is ambiguous without additional information. We tested whether sensitivity to shear motion is enhanced when environmentally relevant depth information is added through binocular disparities. Contrary to previous findings depth information plays no part in shear detection.
List of Figures and Tables

Chapter 1: Introduction

Figure 1.1 Elaborated Reichardt Detector. Localised spatial frequency filters, SF\textsubscript{1} and SF\textsubscript{2}, respond to similar patterns appearing at two different retinal locations. Each response is delayed by the TD units before being multiplied with the other un-delayed signal. If the time delay in TD matches the time taken for the pattern to travel across the retina the correlation of the signals is registered by the multiplication units, X. Units TA provide temporal averaging over a short duration, before the signals are finally subtracted from each other to provide a directional response.

Figure 1.2 The half-height bandwidth of a motion detector with a Gaussian tuning profile. The detector responds maximally when stimulated by a motion at 0 degrees. The height of the horizontal line is positioned at half of this peak. The width of the mechanism is conventionally taken where the horizontal line intersects the Gaussian profile and can be estimated from the x axis by projecting vertically from the intersection points. In this example the half-height bandwidth is from –18 degrees to 18 degrees, i.e. 36 degrees. Alternatively the bandwidth can be computed mathematically by taking the inverse of the Gaussian formula at the intersection points.

Figure 2.1 Construction of a Gabor profile. When a sine wave (blue curve) is multiplied by a Gaussian function (red curve) the resulting windowed sine wave is referred to as a Gabor. Where the sine wave extends infinitely across space in both directions, the
Gaussian window ensures that the resultant waveform is spatially restricted. The spectral composition of the Gabor is limited, being a function of the sine wave and Gaussian profile, jointly.

**Figure 2.2** Construction of a Difference of Gaussian (DoG). When two Gaussian functions (blue and red curves) are subtracted the resulting waveform (solid green curve) is referred to as a Difference of Gaussian. The DoG is spatially restricted, like its component parts, and has a limited spectral content determined jointly by them. See text for details.

**Figure 2.3** A sigmoid function (blue curve) fitted to a hypothetical data set (blue diamonds). The graph shows how the probability of a correct answer (on the y-axis) varies with the signal strength of the independent variable on the x-axis. In this case the independent variable is the contrast of signal in a two alternative, forced choice (2AFC) detection task. Performance range: the choice between 2 options yields a minimum theoretical performance of 0.5 due to guessing, with a maximum performance of 1 at the highest stimulus values.

**Figure 2.4** Sub-asymptotic model fitting. Blue diamonds represent a new data set, similar to that of figure 2.3, except that the observer does not reach 100% performance at the highest stimulus values. Two models have been fitted to this new data set. The blue curve is constrained to reach asymptote at a probability correct of 1. The effect that this has on the 75% threshold estimate (broken lines) is evident by comparison with the red curve, whose asymptotic level was allowed to take on values between 0.96 and 1. The blue model has been distorted, leading to a higher threshold than predicted by the red model.
Figure 2.5  Linear fit. The data from figure 2.3 are fitted here by a straight line. Although the $\chi^2$ error metric has been minimised, the data is not fitted well by this model.

Figure 2.6  Confidence intervals. The broken blue curve was used to estimate a performance threshold at a probability correct of 0.75. The uncertainty attached to this estimate is indicated by a confidence interval (C.I.) that has a defined likelihood of containing the true threshold value. The confidence interval is obtained by fitting two additional models to the data set (green curves). The slope parameter of the green models is identical to that of the best-fitting blue model. Their position along the stimulus axis (x-axis) is determined not by minimisation of the error between the data and the model, but by aiming the optimisation procedure at a defined error value based on the error obtained in fitting the best-fit model. See text for details. The C.I. is bounded by the stimulus values found by projecting from $p(\text{correct})=0.75$ to each of the green models (broken lines).

Chapter 3:  Direction Bandwidths of Rotation, Radial and Translation Motion Mechanisms

Figure 3.1  Schematic of a 2*2AFC detection & discrimination trial. Stimuli were presented in two spatial locations, only one of which contained a field of 200 low contrast Difference of Gaussian dots. In the null location, right, no dots exist. The signal location, left, comprises a population of dots with identical direction component, in this case translating approximately 40 degrees to the right of vertical. Arrows shown here were not present in the
dynamic stimulus and are added to assist the reader. The subjects nominated: a) in which location the signal occurred, and b) whether the signal dots flowed to the right or left of vertical.

Figure 3.2 Response tuning curves of directionally sensitive motion mechanisms. Figure 3.2A shows response activity to high contrast stimuli on the y-axis is plotted as a function of the stimulus direction, and is represented here as a Gaussian function. Mechanism B responds maximally to vertical motion, Mechanisms A & C respond maximally to motion slightly left and slightly right of vertical, respectively. Figure 3.2B shows the reduced response of the same mechanisms to low contrast stimuli, as used in Experiment 1. The mechanisms now do not overlap in their responses, so psychophysical performance is mediated solely by the most responsive mechanism.

Figure 3.3 Psychometric functions showing how the probability of correct detection and discrimination performance (on the y axis) varies with contrast (on the x axis). Separate graphs are shown for each angle of motion presented w.r.t. vertical. At shallow angles of presentation discrimination performance is very poor, not reaching asymptote. As the angle of motion increases discrimination performance improves. Detection performance is good regardless of the direction of the motion.

Figure 3.4 Detection and discrimination thresholds as a function of the direction of motion w.r.t. vertical. Each panel shows detection and discrimination performance for each observer separately under each condition of dot centre spatial frequency. Detection thresholds were fitted with a straight line, discrimination thresholds with a logarithmic function. The equations of these fits
are reproduced on each figure. Points of convergence were calculated by solving the equations simultaneously for a common value, these being taken to represent the point at which responses were independent of the contribution of the vertical motion mechanism. Error bars are 95% confidence intervals.

Table 3.1 Bandwidth estimates of local translation mechanisms in degrees (column 2). Estimates of half of the full width of a directional motion analyser taken from the point of convergence of curves best-fitting detection and discrimination thresholds, see text and figure 3.4. As directions tested were symmetrical about zero degrees, the data were collapsed, removing the sign of the direction. Consequently, full width bandwidths are found by doubling the calculated widths, column 3. Bandwidths are conventionally quoted as full width at half height (column 4), which can be calculated on the assumption that the response profiles of mechanisms are Gaussian, and that their measured full width is equivalent to 6 standard deviations (see text for details).

Figure 3.5 Neighbouring motion detectors with Gaussian response profiles. The horizontal line represents the half-height level traditionally used to characterize mechanism bandwidths. Take curve B to be the response profile of a mechanism sensitive to pure rotation motion. Curve A detects inward-going spirals, and curve C detects outward-going spirals. Assuming that the empirical measure of width reflects the point at which curve B no longer contributes to direction judgements i.e. has fallen to a negligible level (scenario 1), then it can be taken to be equivalent to 6 standard deviations, from which the half-height bandwidth of curve B is easily calculated. Alternatively, under the assumption that the width measured empirically reflects the distance between
the points at which curves A & C intersect with curve B (scenario 2, 'winner-take-all' combination rule), the half-height bandwidth is not calculable without knowledge of the separation between mechanisms.

Figure 3.6  Schematic of a 2*2AFC detection & discrimination trial. Stimuli were presented in two spatial locations, both of which contained a field of 200 Difference of Gaussian dots. In the 'noise' location, right, all dots have a direction component chosen at random to produce incoherent motion signals. The signal location (signal + noise), left, comprises two populations of dots. One population of dots shares an identical direction component, in this case translating approximately 40 degrees to the right of vertical (filled dots). The remaining dots each have a randomly chosen direction component (unfilled). Arrows and filling shown here were not present in the dynamic stimulus and are added to assist the reader. The subjects nominated: a) in which location the signal occurred, and b) whether the signal dots flowed to the right or left of vertical.

Table 3.2  Bandwidth estimates of global translation mechanisms in degrees. Estimates of half of the full width of a directional motion analyser taken from the point of convergence of best-fitting detection and discrimination curves. As directions tested were symmetrical about zero degrees, the data were collapsed, removing the sign of the direction. Consequently, full width bandwiths (column 2) are found by doubling the calculated widths. Bandwidths are conventionally quoted as full width at half-height, which can be calculated on the assumption that the response profiles of mechanisms are Gaussian, and that their measured full width is equivalent to 6 standard deviations (see text for details).
Table 3.3 Summary of the findings of studies where the breadth of the tuning functions of motion sensitive mechanisms have been estimated or implied.

* calculated from the authors' width estimates, which has been assumed to reflect the full width of a single mechanism, whose response profile is presumed to be Gaussian.

Figure 3.7 Schematic of a detection and discrimination trial. Both spatial locations contain a 200 dot RDK. Panel B has random motion as indicated. In Panel A if the dots followed the blue arrows they would form a pure rotation. In fact an angular offset is added to this ideal rotation trajectory such that dots follow the red arrows, simulating an inward spiral motion. Several levels of the angular offset variable were explored, a motion coherence threshold for both detection and discrimination being established for each. Subjects were required to detect in which location the spiral appeared (detection), and whether it was an inward or outward spiral (discrimination).

Figure 3.8 The response function of a hypothetical detector tuned to rotation motion, modelled as a Gaussian function. The blue arrow indicates the stimulation presented by a set of ideal local motions forming a rotating pattern (see text and figure 3.9). The red arrow represents the stimulation caused by a spiral pattern, formed by perturbing the ideal rotation trajectories by an angular offset of -30 degrees. The detector can be said to have a bandwidth along the dimension of angular offset.

Figure 3.9 Rigid rotation. In the diagram a two frame apparent motion stimulus is depicted, simulating a rigid rotation. All three dots
must travel from their start point to their destination in the same
time. As time is constant the length of the arrows is proportional
to the speed of each dot. It is apparent that the speed of each dot
in such a pattern is related to its eccentricity from the centre of
rotation.

Figure 3.10 Schematic of a detection and discrimination trial. Both spatial
locations contain a 200 dot RDK. Panel B has random motion as
indicated. In Panel A if the dots followed the blue arrows they
would form a pure expansion. In fact an angular offset is added to
this ideal rotation trajectory such that dots follow the red arrows,
simulating clockwise spiral motion. Several levels of the angular
offset variable were explored, a motion coherence threshold for
both detection and discrimination being established for each.
Subjects were required to detect in which location the spiral
appeared (detection), and whether it was clockwise or anti-
clockwise (discrimination).

Table 3.4 Bandwidth estimates of global radial and rotation mechanisms in
degrees. Estimates of half of the full width of a directional motion
analyser taken from the point of convergence of best-fitting
detection and discrimination curves, see Experiment 1 for details.
As directions tested were symmetrical about the nominal motion
type (rotation or radial motion), the data were collapsed,
removing the sign of the direction. Consequently, full width
bandwidths are found by doubling the calculated widths, column
2. Bandwidths are conventionally quoted as full width at half
height, which can be calculated on the assumption that the
response profiles of mechanisms are Gaussian, and that their
measured full width is equivalent to 6 standard deviations (see
Experiment 1 for details).
Chapter 4: The Dependency of Sensitivity on Spatial Frequency in Radial Flow

Figure 4.1 Schematic of a 2AFC detection trial. Stimuli were presented in two spatial locations, both of which contained a field of 75 Difference of Gaussian dots. In the 'noise' location, right, all dots have a direction component chosen at random to produce incoherent motion signals. The signal location (signal + noise), left, comprises two populations of dots. One population of dots referred to as 'signal' (filled) are arranged to produce a radially structured pattern, either contraction, or expansion (shown). The remaining dots each have a randomly chosen direction component (unfilled). Arrows and filling shown here were not present in the dynamic stimulus and are added to assist the reader. The subjects nominated the location in which they perceived radial motion.

Figure 4.2 Element size changes of Difference of Gaussian (DoG) elements. A veridical simulation of an approaching object increases in retinal size between first (left) and second (right) samples of its motion. The growth rate of elements was manipulated independently of the rate at which elements diverged from each other. Panel A illustrates zero growth rate between samples. Subsequent panels show veridical growth (B), twice the veridical rate (C) and four times the veridical rate (D).

Figure 4.3a Simulating motion in depth. The schematic shows the geometry of the 'reference' motion in depth simulation, i.e. dots distributed on a fronto-parallel plane are to 'approach' the observer. The dots at point A in the upper panel, at time 1, are to simulate moving to point B, lower panel at time 2. To calculate their image
size and spatial relationships in the original plane (the computer
display), project from the observation point through the new
position. The simulated motion-in-depth has preserved the
relationships between dot size and the centre-to-centre distance
between texture elements (divergence).

Figure 4.3b The geometry of simulating a field of two random dots (black
rectangles) approaching an observer, with dot-size increasing at a
greater rate than specified by divergence. The dots at point A in
the upper panel (at time 1) are to move to point B, lower panel at
time 2. To calculate their spatial relationships and size in the
original plane, project from the observation point. To distort
element growth rate the size of the dots in the projection plane
has been exaggerated. The centre-to-centre relationships between
elements has been preserved, keeping divergence rates among
dots equal to the 'reference' simulation depicted in figure 4.3a.

Table 4.1 Mean change in the peak spatial frequency of RDK dots as a
function of speed and element growth conditions.

Figure 4.4 Coherence thresholds for detecting radial motion. Three curves
are shown, one for each speed condition explored. Thresholds on
the y-axis are plotted as a function of the element growth rate, on
the x-axis. Each data point represents the mean threshold of three
observers. Error bars are one standard error of the mean.

Figure 4.5 Luminance plots in two spatial dimensions (x and y) showing the
effect of filtering a binary dot. Refer to Panel A. Convolving a
dot (left) with a high frequency filter (centre) results in an
isometric dot of defined spatial frequency content. This will be
referred to as a 'circular grating'. Panels B and C show how the
same filter applied to different dots results in spectrally similar dots, differing in size.

**Figure 4.6** Frequency spectra of a sample of the filtered dots of Experiment 2. Dots of different sizes were convolved with the same filter, resulting in isotropic elements differing in size, yet sharing a common spatial frequency composition. Sizes of dots prior to filtering are listed in the insert.

**Figure 4.7** Coherence thresholds for detecting radial motion. Three curves are shown, one for each speed condition explored. Thresholds on the y-axis are plotted as a function of the element growth rate, on the x-axis. Each data point represents the mean threshold of three observers. Error bars are one standard error of the mean.

**Chapter 5: Nulling the Radial Motion Speed Illusion**

**Figure 5.1** Example of the stimuli used in a typical speed matching experiment. An animation is shown in each of two temporal or spatial locations. In one location a field of dots rotates about the centre of a window (left panel). Their direction could be clockwise or counter-clockwise. In the alternative location dots radiate relative to the window’s centre, producing an expanding or contracting pattern. To ensure patterns are equated in speed and direction a radial pattern is created, then a 90 degree directional offset is added to each dots directional parameter to produce rotation. Since the same process is used to generate both patterns, and the 90 degree offset is a constant, both patterns are equated in terms of speed and direction. The observer’s task is to nominate which location appeared the faster.
Figure 5.2  The geometry of crossed (panel A) and uncrossed (panel B) disparities. Point F is the point of fixation point O is the location of an object in the scene. In panel A the object is closer than the fixation plane, and its projection falls on opposite sides of the fixation points on the retinae, as indicated by the directions of the arrows. Panel B shows how an object further from fixation projects to the same side, relative to the fixation image.

Figure 5.3  A stereo pair showing an expanding stimulus. Each image shows a field of dots travelling in the directions indicated by the arrows. By viewing through a Wheatstone stereoscope, only one image is seen by each eye. Each image contains exactly the same spatial arrangement of dots, but those in the right panel are displaced to the right relative to the surrounding black frame. The mark on the top edge of the windows is the point that subjects were asked to fixate during a trial.

Figure 5.4  Results of Experiment 1. The magnitude of the speed illusion is shown for three observers separately (figures 5.4A, B, C). Plotted on the y-axis is the speed of the dots of a rotating pattern (as a factor of the radial speed) that provided a perceptual match to the (fixed speed) radial pattern. The speed illusion is shown as a function of the amount of fixed horizontal disparity on the x-axis, and is shown separately for each of the three observers. The two curves shown on each graph are for crossed and uncrossed disparities, and results for zero disparity are plotted separately on the figure. See legend. The horizontal line shows the dot speed expected if subjects had perceived a match (in retinal speed) terms between rotation and radial patterns.
Figure 5.5  Comparison between the magnitudes of the speed illusion in Experiment 1 (filled bars) and Experiment 2 (open bars). In Experiment 1 the disparity applied to the radial pattern remained at a constant level throughout a trial. In Experiment 2 the disparity varied during the trial, consistent with motion through depth. The data show that the magnitude of the speed illusion was not affected by the dynamic disparity manipulation of Experiment 2.

Figure 5.6  Comparison between the magnitudes of the speed illusion in Experiment 1 (filled bars), Experiment 2 (open bars) and Experiment 3 (grey bars). In Experiment 1 the radial pattern had a fixed disparity throughout each trial. In Experiment 2 the disparity varied during the trial, consistent with motion through depth. In Experiment 3 disparity varied throughout a trial, and the size of dots changed consistent with the motion-in-depth signalled by the disparity change. The pattern of results across all three experiments shows the speed illusion to be insensitive to these depth manipulations.

Chapter 6: Discrimination of Speed in Radial, Rotation and Translation Motion Mechanisms

Figure 6.1  Example of stimuli used in Experiment 1, rotation condition. Two Random Dot Kinematograms (RDK’s), were presented, each containing a global pattern of motion. The observer’s task was to choose which pattern rotated faster. Two other global motion patterns were used, radial (expansion and contraction) and translation. Both animations were carefully matched for average speed, despite the differing rates of rotation, by ‘speed balancing’. See text and figure 6.2 for details.
Figure 6.2  ‘Speed balancing’ in the signal-in-noise stimulus. When two animations containing global motion are presented in the speed discrimination task of Experiment 1, the signal speed of one pattern, the ‘Reference’ is fixed (Panel B). The signal speed in the ‘Test’ pattern is manipulated (Panel A, double headed arrow) to generate a psychometric function. This would normally result in a difference between the average speeds of the test and reference patterns that could be used to discriminate between them. To remove this we manipulated the speed of the ‘noise’ dots in the reference pattern (Panel B, double headed arrows) in concert with the signal speed in the test pattern. This manipulation ensured that both animations shared the same average and peak speeds at all times. Speed profiles are triangular in shape reflecting how dot density varies with eccentricity, i.e. there are more high-speed dots at the outer circumference of a rotating disc than there are near its centre. More dots appear in the periphery, and they travel at greater speed.

Figure 6.3  Examples of psychometric functions relating the probability of seeing the test pattern as faster (y-axis) to the median speed of the test pattern (x-axis). Panels A & B show data for the Radial condition. (NB Observer AS was unable to perform this task). Panels C, D & E show data for the Rotation condition, and panels F, G & H show the Translation data. Blue diamonds signify the empirical data points, the best-fitting cumulative Gaussian (solid purple line) is flanked by curves delimiting the 95% confidence interval.

Figure 6.4  Histogram summaries of the speed discrimination thresholds obtained for each observer. Panel A shows thresholds from the
low-speed condition (median speed=1.57 degrees/s), panel B shows the same information from the high-speed condition (median speed=3.14 degrees/s). Error bars are 95% confidence intervals.

Figure 6.5 Example of stimuli used in Experiment 2. A pair of concentric annuli equated for stimulus area, were presented. Each comprised a field of 200 randomly positioned Difference of Gaussian dots. Both annuli shared the same direction of rotation, clockwise or anti-clockwise, chosen at random on each trial. The observers judged which annulus appeared faster. Three conditions of annulus diameter were used: Panel A ‘Large-Medium’; Panel B ‘Large Small’; Panel C ‘Medium-Small’. The larger diameter annulus was always the fixed-speed reference pattern. See Methods for dimension details.

Figure 6.6 Points of subjective equality (PSE) between concentric rotating annuli of unequal diameter (Experiment 2). For each observer the curves show how the PSE’s (y-axis) vary with the mean speed of the reference pattern (x-axis). The reference was always the larger of the pair. Data (blue diamonds) are fitted with a linear regression, whose equation and R² value are reproduced on the graphs. The solid yellow line is the performance predicted if the subjects were matching pattern speeds using local linear dot speeds. The solid purple line shows the predicted match if angular velocity were being used.

Table 6.1 Speed discrimination thresholds obtained with concentric rotating annuli. Thresholds are expressed as the percentage speed increase required by the test pattern over the reference pattern for the test to be judged faster on 75% of trials. This speed was calculated by
taking the inverse of the best-fitting cumulative Gaussian function to the subject’s psychometric function. Thresholds were only calculated where the subject’s performance on the task was consistent with the use of angular velocity, i.e. when the point of subjective equality across a range of speed conditions qualitatively matched the point of objective equality predicted by an angular velocity matching strategy.

Figure 6.7 Example of stimuli used in Experiment 3. A pair of adjacent annuli that were equated for stimulus area were presented. Each comprised a field of 200 randomly positioned Difference of Gaussian dots. Both annuli shared the same direction of rotation, clockwise or anti-clockwise, chosen at random on each trial. The observers judged which pattern appeared faster. Three conditions of annulus diameter were used: Panel A ‘Large-Medium’; Panel B ‘Large-Small’; Panel C ‘Medium-Small’. The larger diameter annulus was always the fixed-speed reference pattern. See Methods for dimension details.

Figure 6.8 Points of subjective equality (PSE) between non-concentric rotating annuli of unequal diameter (Experiment 3). For each observer the curves show how the PSE’s (y-axis) vary with the mean speed of the reference pattern (x-axis). The reference was always the larger of the pair. Data (blue diamonds) are fitted with a linear regression, whose equation and $R^2$ value are reproduced on the graphs. The solid yellow line is the performance predicted if the subjects were matching pattern speeds using local linear dot speeds. The solid purple line shows the predicted match if angular velocity were being used.
Table 6.2 Speed discrimination thresholds obtained with non-concentric rotating annuli. Thresholds are expressed as the percentage speed increase required by the test pattern over the reference pattern for the test to be judged faster on 75% of trials. This speed was calculated by taking the inverse of the best-fitting cumulative Gaussian function to the subject's psychometric function. Thresholds were only calculated where the subject's performance on the task was consistent with the use of angular velocity, i.e. when the point of subjective equality across a range of speed conditions qualitatively matched the point of objective equality predicted by an angular velocity matching strategy.

Figure 6.9 Summary histogram of speed discrimination thresholds as a function of observer and viewing configuration that resulted in matching using angular velocity.

Chapter 7: Do Relative Motion and Binocular Disparity Interact in the Detection of Shear Motion?

Figure 7.1 As an observer travels in the direction indicated by arrow A, a typical landscape gives rise to opposing directions of motion on the retina, arrows B and C. When combined with depth information this relative motion (RM) is diagnostic of the observer's direction of travel. A scene containing a smooth change in depth produces a smooth speed gradient, with velocity falling to zero at the fixation point.

Figure 7.2 The retinal velocity gradient arising from the ego-motion depicted in figure 7.1. Arrows indicate the direction of motion, with the sign of the motion in the lower hemi-field being the inverse of that in the upper hemi-field. The length of each arrow
is proportional to the speed component of the motion vector. The progressive change in arrow length with vertical displacement from the centre of fixation (black dot) reflects the smooth speed gradient observed when viewing a flat ground plane.

Figure 7.3 Schematic representation of the three conditions of depth structure simulated using binocular disparity. Panel A shows the RDK elements receding smoothly in depth, as a landscape would (‘slanted’). Panel B shows dots randomly distributed in depth (random’) and Panel C shows all dots at a constant depth (‘fixed’). Distances shown were calculated with respect to the screen of the VDU, which was viewed from a distance of 57cm. Filled dots are signal, whose directions produce shear motion (not shown), unfilled dots are directionally random noise dots.

Figure 7.4 Motion coherence thresholds for the detection of shear motion (vertical axis) as a function of median speed (horizontal axis) and disparity condition (red, blue and yellow curves). The first three graphs show data for each of the three observers separately. Error bars are 95% confidence intervals. As the data for each subject were similar the final graph shows the data for all three observers collapsed. Error bars are one standard error of the mean.

Figure 7.5 Lifetime=6 frames. Motion coherence thresholds for the detection of shear motion (vertical axis) as a function of median speed (horizontal axis) and disparity condition (red, blue and yellow curves). As the data for each subject were similar the graph shows the data for all three observers collapsed. Error bars are one standard error of the mean.
Figure 7.6  Motion coherence thresholds for the detection of shear motion (vertical axis) as a function of median speed (horizontal axis) and disparity condition (red, blue and yellow curves). The three graphs show data for each of the observers separately. Error bars are 95% confidence intervals.

Figure 7.7  Schematic representation of the three conditions of depth structure simulated in Experiment 3. Signal dots (filled) obey the three disparity conditions, as before. Noise dots now occupy their own depth plane at all times.

Figure 7.8  Motion coherence thresholds for the discrimination of shear direction. Thresholds are shown on the y-axis, as a function of disparity type on the x-axis, for each of the three observers separately.
Chapter 1
Introduction

1.1 Abstract
This introductory chapter covers topics crucial to the understanding of the work presented in the thesis. The opening section presents a review of the existing literature pertaining to the analysis of optic flow, as documented by researchers working in psychophysics, electro-physiology and computational modelling.

The second section deals with technical aspects of the experimental work reported here. Four topics are addressed, including a description of the methods used to sample the observer's psychometric function; stimulus design; processing techniques used to extract key parameters of human performance from the raw data produced by experimentation. Finally, the use of personal computer equipment, in terms of both hardware and software is considered.
1.2 Literature Review

1.2.1 The Utility of Motion Information

A note on terminology. Unlike electrophysiologists, psychophysicists can never be sure whether a single neuron or a group of neurons is responsible for the patterns of performance shown by observers. In common with the psychophysics literature hypothetical ‘mechanisms’ (Graham, 1985) or an appropriate synonym such as ‘detectors’ or ‘analysers’ will be referred to in order to describe the underlying neural hardware that is presumed to be responsible for sensing visual stimuli.

The ability to process image motion is exploited in many ways by living organisms. The study of this ability has value to those interested in how human and animal brains function, and to those concerned with automating processes that previously required human involvement. Nakayama (1985) provided a comprehensive review of the tasks that motion detection helps us to perform. These include: navigation through complicated environments; estimating time to collision during navigation; the detection of moving objects; the control of eye movements; image segmentation, and the analysis of the relative depth structure of the environment. This thesis is concerned with the motion mechanisms that detect and interpret the retinal motion produced by self-movement in the world. These signals were termed ‘optic flow’ by Gibson (1950, 1979), who recognised that much valuable information was contained in the complicated array of motion vectors produced by locomotion.

To understand the work presented here, three different themes will be used to organise the literature. Firstly, theoretical treatments have revealed how motion parameters might be calculated, and how these results can be used to perform real world tasks like the computation of heading direction. Secondly, the physiological investigation of the visual system of the primate animal model has suggested biologically plausible mechanisms for the analysis of complex retinal
motion. Finally, fundamental to any movement are the attributes of speed and direction. Much psychophysical effort has been directed at investigating the ability of humans to code speed and direction, and the results of this work are described below. Naturally, the results from each field influence one another, and their interaction guides the direction of future research.

Unfortunately, two conventions have arisen in the literature on complex patterns of motion, where common terms are used having different meanings. Some investigators refer to translation to describe the motion of the observer, where others talk about the translation of retinal signals. Similarly, rotation is sometimes used in body coordinates, e.g. rotation of the head or eye, versus a rotating field of motion signals on the retina. To avoid confusion all references to translation, rotation (and radiation, which has not been duplicated) throughout the thesis will refer to retinal motion signals, not observer motion.

1.2.2 The Computational Approach

The matching or correspondence model.

Object movement is a change in spatial location over time. A model seeking to sense object motion could identify its locations at different moments in time calculating direction from the location information and speed from the distance travelled divided by the time taken. This scheme faces a major hurdle in ensuring that the correct object is identified between samples of the scene. This difficulty has been termed the ‘correspondence problem’ and has been addressed by the matching model of Ullman (1979). To solve it the algorithm of Ullman’s motion detector makes similarity comparisons between candidate correspondences and matches candidates according to a globally based cost/benefit analysis. While it is plausible that high-level forms of motion detection exist, there are reasons to seek out low-level mechanisms. Matching models have been criticised on the grounds that highly textured images, common in our everyday experience, are likely to present too many potential correspondences for a matching model to perform adequately (Heeger, 1987).
Furthermore the model takes as its starting point some of the very properties we would like to end up with (object properties), and which motion can be so helpful in extracting. The well-known phenomena of structure from motion (e.g. Longuet-Higgins & Prazdny, 1980) illustrates that motion may not only precede but also be necessary for the segmentation of a scene into constituent objects. For these reasons we know that motion is coded very early in the visual pathways, and since low-level cells do not have object features available to them, models working on quite different principles have been proposed.

Correlator models.
The movement of an object gives rise to a change in the luminance distribution over time. Rather than identify the object and register its locations over time, a class of 'low-level' models operating more directly on the changing luminance distribution has been developed. The early work of Reichardt (1969) on the vision of insects gave rise to the 'correlator' design. In its elementary form the model samples the scene at two spatial locations. Movement occurring between these points will cause a pattern to appear at one location, then a short time later at the other. The signals from each spatial location are brought together, with a time delay applied to the first. If this delay matches the time taken for the pattern to change spatial locations the signals will coincide, indicating movement. A pair of such units (illustrated by the Elaborated Reichardt Detector of van Santen & Sperling (1985) at figure 1.1) is capable of producing directional information through the relative magnitude of opponent signals. Models from Adelson & Bergen (1985) and Watson & Ahumada (1985) cast their similar mechanisms in terms of spatio-temporal filters, replacing the notions of distinct spatial location with spatial phase, and time delay with temporal phase. They do however retain the basic opponent motion calculation. The presence of spatial filters at the input to these models gives rise to the need for many such units, each covering a different spatial range, suitable for detecting movement over both small and large distances.
Figure 1.1 Elaborated Reichardt Detector. Localised spatial frequency filters, SF₁ and SF₂, respond to similar patterns appearing at two different retinal locations. Each response is delayed by the TD units before being multiplied with the other un-delayed signal. If the time delay in TD matches the time taken for the pattern to travel across the retina the correlation of the signals is registered by the multiplication units, X. Units TA provide temporal averaging over a short duration, before the signals are finally subtracted from each other to provide a directional response.
The single channel, pre-filter model.

A model that inhabits a middle ground between low-level spatial frequency matching on one hand and global feature matching on the other is that of Morgan (1992) (see also Morgan & Fahle, 1992 and Morgan & Mather, 1994). The model of Morgan and colleagues differs from low-level models in the filtering applied to the input. As pointed out in the previous section, the generally accepted existence of several 'independent' spatial frequency channels has led to the assumption that motion detection occurs separately in spatial frequency bands, before these motion signals are combined to produce a final decision on the likely motion parameters present. (The initial conception that spatial frequency channels are independent has not survived intact. See Wilson & Wilkinson, 1997). Their single-channel, pre-filter model places just one filter, of Gaussian profile according to later work (Morgan & Mather, 1994), ahead of motion detection. Watt & Morgan's (1985)'zero-bounded regions' are the features that are matched, the zero boundary being the zero crossing of Marr & Hildreth (1980) that is taken to signify an object's edge. If such a bounded region can be taken as the outline of an object then the model detects the motion of high-level features. On the other hand its earliest incarnation used the zero-crossings themselves as the matched features, without the further constraint that they should form a bounded region. This ambivalence regarding the exact nature of the representations to be used in the matching process stems from the fact that the model was developed primarily to account for the empirical data on $D_{\text{max}}$. The name $D_{\text{max}}$ is given to the maximum distance across which a dot can be displaced while still eliciting some measure of apparent motion in the observer, for example direction discrimination.

The Gradient Model.

Following on from the work of Marr & Hildreth, (1981) in the static spatial domain, where object boundaries were identified with high rates of luminance change (a luminance gradient), Marr & Ullman (1981), applied the luminance gradient idea to the space-time domain, suggesting that identifying and
comparing steep luminance and temporal gradients might yield the speed and direction estimates required (see also Harris, 1986; Mather, 1990). This technique has the advantage of using a very low-level means of extracting one simple feature from the array, achievable using the properties of low-level cells, and without requiring an entire object level representation. This scheme has had its adherents despite the popularity of later models working on different principles, and despite its unreliable performance at peaks and troughs of luminance. This difficulty arises because the gradient detector divides the derivative of intensity over time with the derivative of intensity over space. Zero or near zero values in the spatial component result in very large quotients, leading to unreliable parameter estimates (Johnston & Clifford, 1994). Heeger (1987) has pointed out that highly textured images present great difficulties for models working on the gradient principle, as the gradient direction may vary randomly between points. These models are held to be prone to producing errors if noise is introduced to the scene, or the image intensity is not preserved between frames of an apparent motion sequence (Perrone, 1990).

1.2.3 Modelling Optic Flow Analysis
While the foregoing algorithms provide an account of local motion detection, they do not serve as general models: they are restricted to one direction of motion, and are confounded by the aperture problem. The aperture problem refers to the inability of a cell restricted in the spatial extent of its analysis to signal more than one direction of motion. A stimulus extending beyond this cell’s receptive field (rf) may move obliquely through the rf, so speed and direction are confounded in the cell’s output. To account for the perception of motion in two dimensions many have proposed that local motion signals are ‘pooled’. Combining the outputs of local movement detectors gives the ability to register the type of complicated wide-field patterns encountered when moving through a scene comprising static as well as dynamic objects. The optic flow that such relative motion generates has been the subject of much theoretical and practical modelling. This interest reflects the important role that may be played
by optic flow in visual navigation (Rieger & Toet, 1985; Warren & Hannon, 1990), detecting object boundaries (Hildreth, 1992), coding the movement of independent objects in the field (Zemel & Sejnowski, 1998), and generating opto-kinetic eye-movements and smooth pursuit (Dursteler and Wurtz, 1988; Komatsu and Wurtz, 1988; Erickson and Dow, 1989; Kawano et. al., 1994). Models of optic flow processing are most often identified with extra-striate area medial superior temporal area (MST) in primates, where large receptive field sizes and sensitivity to complex, optic flow-like stimuli are found (e.g. Saito et. al., 1986; Tanaka et. al., 1986; Lagae et. al., 1994).

In a recent review, Perrone (2001) suggested a distinction between biologically-inspired ‘template models’ of optic flow processing and the more abstract mathematical ‘vector models’. The template class of models are so named for their motion matching architecture. In order to register a complex pattern of motion a hypothetical detector receives signals from multiple local velocity detectors, specifically organised in their spatial distribution to match a typical or commonly occurring pattern. Different patterns demand different organisational structures, but each is a simple prototype awaiting an appropriate trigger. Such a technique has the advantage of being able to side-step complex tasks (Glunder, 1990). For instance, computing speed could be achieved by estimating distance travelled and time taken, followed by the appropriate division: distance/time. However, a template model responds when its spatial and temporal parameters are met, without intermediate representations of distance and time being computed.

Vector models make use of the observation that a flow field can be regarded as a set of basic flow components (Koenderink & van Doorn, 1975). This mathematical fact has led to many studies investigating how optic flow patterns could be decomposed into elementary constituents for use in the types of tasks that mobile organisms need to perform. However, the attempt to retain biological plausibility has been reflected in the approach to one testing problem. In an organism with the ability to move the eyes and head, as well as the body, retinal signals confound all three of these sources of motion. Vector based approaches
generally seek to use just the retinal motion information to unconfound these signals (Rieger & Lawton, 1985; Koenderink & van Doorn, 1981; Heeger & Jepson, 1992), proving that such signals are sufficiently informative. However, psychophysical evidence (e.g. Royden, Crowell & Banks, 1994) suggests that extra-retinal signals may be used in determining heading direction accurately by accounting for self-generated sources of retinal motion such as eye rotation, and template modellers have recently begun to incorporate extra-retinal signals to assist in segregating the various signals (Beintema, 1998).

Physically instantiated models can be trained to make use of simulated optic flow in performing a task (Hatsopolous & Warren, 1991; Lappe et al., 1996). For example Lappe et al., (1996) trained a neural network to perform heading identification when presented with a variety of realistic optic flow patterns. Once trained, such models can be examined to see what kind of representations have been encoded in learning to make use of the information available. Many authors claim that synthetic neurone receptive field properties mimic those of real neurones found in electro-physiological investigations of animal models. Examples include the relative frequency of units selective for expansion, contraction, rotation and translation (model: Zemel & Sejnowski, 1998; physiology: Graziano et al., 1994); the changing response of individual neurones, depending on the position of the focus of expansion (FOE), (model: Lappe et al., 1996; physiology: Duffy & Wurtz, 1995); the occurrence of position invariant responses in neurones (insensitivity to the position of the FOE) (model: Zemel & Sejnowski, 1998, physiology: Graziano et al., 1994); and lack of immunity to superpositions of non-preferred and preferred stimuli (model: Perrone & Stone, 1998; physiology: Orban et al., 1992). This last point bears on a concept common to both vector and template models—that of decomposition. It is often presumed that an efficient way to deal with the complexity of multiple motion vectors is to decompose them into orthogonal constituents, analogous to the way that the statistical technique Principal Components Analysis does to complex data sets. In template models this decomposition is often referred to as a basis set, or canonical representation, and
the templates responsible for achieving it are usually radial, rotation, deformation and translation detectors. These detectors have been linked to the theoretical components identified by Koenderink & van Doorn (1975), namely, \( \text{div}: \) the rate of expansion, \( \text{curl}: \) the rate of rotation. A further category, \( \text{def}: \) the rate of deformation, has received a little physiological support from Orban et. al., (1992), and some behavioural support from Meese and Harris (2001a, 2001b). Neurones whose response selectivities are immune to contamination by multiple components (e.g. rotation and radiation combined) would be consistent with decomposition to a basis set. However, physiological studies have recorded many departures from this ‘ideal’ representation. Receptive field selectivities that form a continuum of responses from radial through spiral to rotation patterns have been found (Duffy & Wurtz, 1991), as have rf’s responsive to simultaneously presented combinations of canonical components (Graziano et. al., 1994). There is also some behavioural evidence against the decomposition hypothesis. Duffy & Wurtz (1993) and Grigo & Lappe (1998) found that combining radial expansion with translation leads to a misperception of the centre of expansion, implying lack of the ability to separate the components effectively. This effect was modelled by Lappe & Duffy (1999), showing that a population of neurones was capable of emulating the behavioural evidence without resorting to basis set selectivities. The emerging view is that MST (and other extra-striate areas) form a complex set of representations of patterns of optic flow, depending on their utility and their frequency of occurrence (Lappe et. al., 1996; Irvins et. al., 1999).

1.2.4 Physiology
Electro-physiologists are most concerned with the biological architecture of the visual system, asking: what biologically plausible structure could be responsible for coding optic flow? One influential proposal has global motion detection accomplished in multiple stages of processing (Van Essen & Maunsell, 1983). Early motion-sensitive cells in V1 have spatially restricted receptive fields, and
so are ill suited to respond unambiguously to the movement of objects that exceed their field of view (the aperture problem, described above).

Anatomical investigation of the primate visual architecture, primarily in the macaque animal model, has revealed parallel pathways leading from the retinae to the cortex via the lateral geniculate nucleus (LGN). These pathways are referred to as the M and P pathways, and account for approximately 90% of the projections from the retinae (Shapley & Perry, 1986). Sub-cortical ganglion cells, magnocellular and parvocellular neurons, while generally responsive to temporal modulation when a stimulus is moved through their receptive field, do not show strong directional selectivity. Cortical area V1 is the site at which this property, so essential for higher order motion processing, is first encountered (Hubel & Wiesel, 1968; Colby et al., 1993; De Valois, Yund & Hepler, 1982). Direction sensitive cells are also found in extra-striate visual areas such as V2 (Zeki, 1978), V3 (Felleman & van Essen, 1987), and V4 (Desimone & Schein, 1987; Ferrera et al., 1993). However, by far the most important areas so far identified for the processing of visual motion are the middle temporal area (MT, also known as V5) and the medial MST. In both of these areas a high proportion of cells exhibit directional selectivity (Albright et. al., 1984; Maunsell & Van Essen, 1983; Desimone & Ungerleider, 1986; Tanaka et. al., 1986). Newsome, Britten & Movshon (1989) demonstrated a correlation between the action of MT cells and the ability of a monkey to discriminate between opposing directions of movement. Stimulating MT cells has been shown to modulate the decision of a monkey when asked to indicate the direction of movement that it perceived in a motion display (Salzman et al., 1990). The M-pathway has been identified as providing the major input to MT by Maunsell et. al., (1990). Maunsell and colleagues selectively blocked the activity of both M and P processing streams, finding that the response of MT neurons was substantially reduced when the M pathway was temporarily lesioned, but remained largely unaffected by P pathway blockade. While MT cells are selective for direction of motion, and are broadly tuned for speed (Maunsell & van Essen, 1983; Lagae et al., 1993), they are insensitive to either the shape or colour of the stimulus (Zeki, 1974). As in
earlier visual areas the receptive field size of MT neurons varies with eccentricity, becoming larger with increasing distance from the fovea (Albright & Desimone, 1987). MT projects to MST where receptive fields are larger in size, and are thought to be especially suited to the detection of wide-field movements (Tanaka, 1998). Such patterns of flow are those typically caused by the motion of the observer through the environment, and may take the form of a rigid translation, expansion/contraction or rotation. Cells selectively responsive to these prototypical patterns are found in dorsal MST (Sakata et al, 1985; Sakata et. al., 1986; Sakata et al., 1994; Saito et al., 1986; Duffy & Wurtz, 1991; Orban et. al, 1992; Graziano et. al., 1994), presumably allowing for the coding of ego motion, especially when one unusual property of these cells is considered. ‘Position invariance’ refers to the lack of sensitivity of radial and rotary cells to the locus of their preferred pattern of motion. The centre of expansion/contraction or rotation fields does not need to be placed at the centre of the receptive field in order to elicit strong cell responses. This may give these cells the ability to code the motion of the animal’s body independently of the gaze direction. Cells in ventral MST exhibit a different pattern of responses to those in the dorsal region, one that seems to be sensitive to the relationship between figure and ground. A large proportion of them register the presence of small stationary stimuli, with the sign of their directional selectivity being contingent on the presence or absence of a background wide field motion (Sugita & Tanaka, 1991; Tanaka et al., 1993). This pattern of responses gives these neurons the property of detecting the relative motion between object and background (Tanaka, 1998).

1.2.5 Psychophysics

When something moves, two crucial parameters of interest are its direction of movement, and its speed. The ability to register both of these attributes has been studied many times. The presence of cells in primate extra-striate cortex that are responsive to wide-field, or global patterns of motion has led to renewed interest in these topics. The medial temporal (MT) and medial superior temporal (MST)
areas contain motion-sensitive cells with much larger receptive fields than those of V1, which projects to them (Maunsell & Newsome, 1987). If these mechanisms represent a further stage of motion analysis, their speed and directional tuning competence may be different to that measured by previous investigators. The following section deals with what is currently known about speed and direction capabilities, as measured by behavioural techniques.

1.2.5.1 Speed and Direction in Local Mechanisms

In the first instance the visual scene is analysed for the presence of motion in discrete regions. This ‘local’ analysis introduces a tension between competing requirements if estimates of both speed and direction parameters are required from the same mechanism. To sense motion it is necessary for a system to somehow register the time taken to traverse a distance. The ability to obtain an accurate local estimate of speed depends on keeping the distance and time values in the equation low, avoiding averaging. Yet the accuracy of a direction estimate is improved if these values are increased (Grzywacz et al., 1994). Calculations of the error involved in local speed discrimination by human observers range from 30% to 100%, indicating that speed is not initially computed with great accuracy, (Vaina et al., 1990; Bravo & Watamaniuk, 1991). When local signals are combined, however, precision increases greatly, with discrimination errors as low as 5% being reported (McKee et al., 1986; Vaina et al., 1990; Bravo & Watamaniuk, 1991). These estimates were obtained using random dot displays with isotropic (circular) elements, so it is certain that subjects were relying solely on their perception of motion direction, rather than using an artifactual orientation cue. The high errors in local motion speed estimation might lead us to conclude that accurate local speed estimates have been sacrificed in favour of directional tuning.

Yet estimates of the directional bandwidths of motion mechanisms indicates that these too are broad, in the region of 130 degrees, (Ball & Sekuler, 1980; Ball et al., 1983). However, there are two problems with previous estimates of directional tuning. Reported values for bandwidth have been estimated from
experiments which have been designed to gauge how great the angular distance between signals is required to be in order that the mechanisms responding to them do not interact with each other. (e.g. Levinson & Sekuler, 1979; Mather & Moulden, 1980; Ball & Sekuler, 1979; Ball & Sekuler, 1980; Ball & Sekuler, 1981). These experiments found that directions of motion had to be between 120 degrees and 150 degrees apart for independence between mechanisms to be evident. These figures have then been interpreted directly as bandwidths of directional tuning. However, a principled estimate of bandwidth is obtained at the half height of the mechanisms response curve. Consider figure 1.2 which shows a Gaussian profile representing the response profile of a motion detector sensitive to a finite range of directions. Conventionally the range of directions to which it responds is quoted at half its maximum response, hence ‘half-height bandwidth’. This definition is useful as an attempt to measure its width at the base would be problematic due to the theoretically infinite nature of the decay of the Gaussian function. In addition, low-level noise would be less likely to contaminate width estimates made at half height. Estimates made for half-height bandwidth have given rise to much smaller figures of 70-90 degrees (Raymond, 1993). The point of independence between mechanisms is a region of the response curves that may slope very gently, and give rise to misleadingly large and variable estimates. Secondly, the design of the random dot stimuli used in some of these experiments did not preclude subjects from integrating across many elements, hence possibly producing bandwidth estimates of global direction detection mechanisms. Directional tuning, as measured by unit recording techniques in monkey, results in bandwidths of circa 40 degrees for V1 cells (De Valois, Yund & Hepler, 1982), and circa 80 degrees for MT cells (Felleman & Kaas, 1983).
Figure 1.2. The half-height bandwidth of a motion detector with a Gaussian tuning profile. The detector responds maximally when stimulated by a motion at 0 degrees. The horizontal line is positioned at half this peak. The width of the mechanism is conventionally taken where the horizontal line intersects the Gaussian profile and can be estimated from the x axis by projecting vertically from the intersection points. In this example the half-height bandwidth is from -18 degrees to 18 degrees, i.e. 36 degrees. Alternatively the bandwidth can be computed mathematically by taking the inverse of the Gaussian formula at the intersection points.

1.2.5.2 Speed and Direction in Global Mechanisms

Measuring the directional tuning bandwidths of complex motion detectors requires that stimuli oblige judgements to be made using global detectors rather than local ones. Evidence has been accumulating from psychophysics (Regan & Beverley, 1978; Morrone, Burr & Vaina, 1995; Burr, Morrone & Vaina, 1998; Snowden & Milne, 1997; Meese & Harris, 2000) and electro-physiology (Tanaka & Saito, 1989; Duffy & Wurtz, 1991; Graziano, Andersen & Snowden, 1994) supporting the existence of mechanisms responsive to complex motions
such as rotation and radial patterns, as well as intermediate spirals (Meese & Anderson, 2002; Snowden & Milne, 1996). These mechanisms have also been characterised as broadly tuned in their directional sensitivity in several psychophysical studies. For example, measuring motion coherence thresholds Morrone et. al., (1999) found enhanced sensitivity to radial and rotation patterns over intermediate spirals. While not ruling out the possible existence of comparatively less sensitive spiral mechanisms, they chose to interpret their results in terms of broadly tuned cardinal detectors. Where evidence is presented for cardinal or basis set components, broad direction bandwidths are often assumed, as these mechanisms need to detect components at intermediate angles of flow in order for them to be coded in conjunction with another detector.

Meese & Harris (2001) showed that summation between individually sub-threshold stimuli declined when their directions exceeded approximately 30 degrees, falling to asymptote at around 90 degrees of separation. Modelling these data, Meese & Harris (2001) proposed that mechanisms of directional tuning bandwidth 80 to 120 degrees provided an account of the results, dependent on the number of mechanisms employed. Similarly, Meese & Anderson (2002) settled on 92 degrees direction bandwidths in a continuum of rotation, spiral and radial detectors to model sub-threshold summation between those patterns of optic flow. Burr, Badcock & Ross (2001) also exploited a summation technique to provide evidence for cardinal optic flow detectors. Observers were more sensitive in discriminating the direction of spatially interleaved orthogonal optic flow patterns when the orthogonal components were spirals than when they were rotation and radial patterns. This result was taken to indicate that radial and rotation motions are relatively independent, whereas spirals share encoding components, therefore enjoying increased sensitivity due to summation.

However, even when intermediate spiral detectors are found, their directional tuning can be said to be broad. One such study is that of Snowden & Milne (1996) who adapted observers to rotation, radial and spiral patterns, separately. Motion coherence detection thresholds were elevated after adaptation. The size
of the effect was contingent on the similarity between adapt and test patterns, being greatest when adapt and test were identical. These results were modelled by mechanisms sensitive to a range of pattern detectors that spanned radial, spiral and rotation sensitive units, their Gaussian mechanisms having standard deviation of 40 degrees, equating to 94 degrees bandwidth at half-height. Although the bandwidths of motion sensitive detectors are often characterised as broadly tuned, i.e. sensitive to a large range of directions, this does not necessarily imply that precise coding of direction is unattainable. While one such mechanism might be unable to identify its input precisely, a population of overlapping mechanisms could do. The terms population, vector and course coding are often used to describe this method. In the colour domain the principle of univariance describes the loss of specific wavelength information when a single broadly tuned mechanism (Long, Medium or Short retinal cone type) transduces its input to a train of neuronal pulses. The activity of any single cone confounds both intensity and wavelength in its output, hence different combinations of frequency and intensity produce the same activity. An equivalent to this in the motion domain are the attributes of luminance contrast and direction of motion. It is therefore possible to obtain the same output from a detector with each of the following combinations of stimulus: 1) an optimal directional signal at a low contrast level, 2) a sub-optimal directional signal but at a high contrast level. A population of overlapping mechanisms would have more than one mechanism responding to the same input, and the ratio of their responses would specify the input signal more precisely for a given number of neurones (Churchland & Sejnowski, 1992. See Pouget et. al., 2000 for a discussion of other attractive features of population codes). Therefore a design incorporating broadly tuned mechanisms would also feature closely spaced mechanisms, such that several adjacent mechanisms would be stimulated by the same input. However mechanisms that are too close would not produce an appreciable change in their population code for small increments of the attribute of interest. It is clear from this analysis that a trade-off could be made between bandwidth and mechanism separation. The issue of bandwidth and separation
will be addressed further in the first experimental chapter, "Direction Bandwidths of Rotation, Radial and Translation Motion Mechanisms".

*Speed*

Smith, Snowden & Milne (1994) found that a wide-field translating pattern needed a velocity increase of circa 10-12% to be distinguished from a reference pattern moving at 4 degrees/s. Extending earlier work on the speed discrimination of simple linear motions, Verghese & Stone (1995) asked observers to discriminate on the basis of speed between two animations, each comprising multiple Gabor patches. They found that speed discrimination thresholds decreased with increasing number of patches (1 patch = circa 19%; 6 patches = circa 11%, interpreted from their figure 3), a finding they attributed to the parsing of visual space into discrete entities. When the area of a single patch was increased in line with that shared across multiple patches thresholds were constant, at approximately 20% (interpreted from their figure 5a).

While global translation may be thought of as an extension of local translation, with no additional interpretation being placed on local speed signals, rotation and radial motions are different. Both radial and rotating patterns may be described in speed terms that are quite different to that of the local speeds that make them up. A rotating object has its speed of rotation expressed in degrees per second (hence often referred to as angular speed) that could in theory be sensed by the visual system. For example a complete rotation of a ball about its centre in one second would have an angular speed of 360 degrees/second. Any texture element at the circumference would have the same angular speed as one near the centre, yet it travels a much greater distance, and hence has a higher speed in terms of distance/time. Assigning the rotating object one value of rotational speed, rather than multiple values of local speed, would seem to be more desirable.
Similarly, radial patterns contain multiple local speeds produced by relative motion between the eye and the scene, their precise values being partly dependent on the structure of the environment. However, the speed of the relative motion is more important to the organism than individual retinal speeds.

To investigate whether the angular speed of rotating patterns is sensed, a series of papers by Werkhoven & Koenderink (1990, 1991, 1993) examined the perception of speed in rotating RDK patterns. In their first study, one experimental condition examined speed discrimination between two temporally separated rotating circles of dots, having equal radius. Observers discriminated between a reference pattern of constant speed (80 degrees/second) and a speed-varying test pattern, showing that lowest difference thresholds of 8-10% were achievable. (Discrimination threshold was taken at the point where observers judged the test pattern as faster on 84% of trials).

In their second paper (Werkhoven & Koenderink, 1991) similar stimuli were used in four experiments examining the dependency of speed discrimination on several factors:

Temporal factors. A range of frame durations (14..114ms) and stimulus durations (57..1829ms) were explored. Long stimulus and short frame durations gave rise to lowest discrimination thresholds of over 5%, though short stimulus and long frame durations compromised thresholds severely, e.g. 60%.

Textural factors. The number of dots (1..32) and dot density were varied to explore how these parameters influenced discrimination performance. (Dot density was manipulated by scattering the dots at different eccentricities around the nominal radius). Discrimination thresholds were found to be lowest (circa 10%) when dots were more numerous, and when they were more densely organised on the nominal radius.

Spatial scale. Both comparison stimuli were presented at a range of viewing distances (7..1200cm) to manipulate their retinal sizes (test and reference always the same size). The results show that discrimination thresholds vary with the
(retinal) radii of the annuli in a u-shaped pattern, i.e. higher thresholds (10..18%) for very long and very short viewing distances, but circa 8-10% for intermediate viewing distances.

Finally, the dependence of discrimination threshold on rotation (angular) speed was explored over the range 10..320 degrees/second to ascertain whether Weber's law held for rotational motion in the way it does for linear motion. For speeds above 80 degrees/second it did, being consistently below 10%, though rates of rotation below 80 degrees/second resulted in progressively more elevated thresholds of 25% and above.

Their third paper (Werkhoven & Koenderink, 1993) explored how speed discrimination thresholds were influenced by differences in the radii of test and reference patterns. In these experiments the reference and test annuli comprised eight dots rotating around a fixed point, again presented in two temporal intervals. Speed discrimination was measured between reference and test patterns over a range of reference speeds (40..160 degrees/second) and radii (the test pattern radius was a multiple of the reference pattern radius, range: 0.25..2). Discrimination thresholds were found to vary with the degree of mismatch between the radii of comparison stimuli. Minimum threshold was obtained when sizes were equal (7%), increasing to more than 20% when radii differed by a factor of two.

As a result of these experiments the authors concluded that local mechanisms were fundamentally responsible for the pattern of speed discrimination performance observed, although their results did not unequivocally support reliance on either local or global mechanisms.

There is evidence that the perceived speed of radial optic flow patterns is not simply that of local speeds making up the pattern. Several investigators have confirmed that radial patterns apparently travel at greater speed than their physical speed would suggest. By matching the perceived speed of a rotation pattern to that of a radial pattern, a bias is found whereby a perceptual match is only achieved when the components of the rotation are travelling faster than
those in the radial pattern. (Bex & Makous, 1997; Bex Metha & Makous, 1998; Clifford, Beardsley & Vaina, 1999; Clifford, Arnold & Wenderoth, 2000; Geesaman & Qian, 1996; Geesaman & Qian, 1998). This effect is known to be caused by the spatial relationship amongst the motion signals making up the radial pattern as the illusory speed is obtained even when comparison rotation stimuli are composed of an identical set of motion vectors, albeit with the spatial arrangement necessary to produce rotation. One explanation for the exaggerated perception of speed is that radial patterns are interpreted as relative motion between observer and environment, and the speed perceived is that of motion-in-depth, rather than of the velocity of component retinal signals. This possibility suggests one reason why the magnitude of the illusion varies so greatly between studies. Estimating speed in depth is problematic without an estimate of distance. Retinally, the same stimulation could be expected of a nearby object moving slowly towards the observer, or a distant object moving quickly.

Mammalian vision incorporates a plethora of cell types, each of which is specialised along several dimensions (Graham 1985; 1989). A comprehensive list would be long, however some important examples of such dimensions include motion, orientation, colour and contrast.

As described in the foregoing paragraphs evidence has been accumulating for the existence of specialised mechanisms that encode complex patterns of motion. Motion constitutes only one dimension that they may respond to, and only by characterising these other domains can a full understanding of their role in visual processing be gained. It was the aim of the work described in the remainder of this thesis to explore the sensitivities of complex motion mechanisms to the following attributes:

Direction of motion. Typically, a motion-sensitive mechanism responds to a limited number of directions. While direction bandwidths of simple translation has been estimated many times, less work has focussed on complex patterns of
motion. In chapter 3 the width of such a range is measured, establishing the bandwidth of detectors sensitive to radial motion, rotational motion and simple translation.

Spatial frequency. The initial analysis of the dynamic visual scene is thought to be carried out by cells that respond to a limited range of spatial frequencies. Little work has explored whether mechanisms that code complex patterns of motion are similarly tuned. The specificity of radial motion mechanisms for spatial frequency is explored in chapter 4.

Speed. Radially expanding patterns have been found to generate a sense of speed in observers that is greater than would be expected of the retinal speeds presented. It has been suggested that observers interpret a radially expanding pattern as motion in depth, and that the speed they sense is that of motion in the third dimension. Chapter 5 investigates this hypothesis to establish whether mechanisms sensitive to radial flow are tuned to motion in depth.

Speed. The capacity to estimate speed accurately is revealed in the ability to discriminate between objects moving at similar speeds. This ability is explored in chapter 6 to establish the speed discrimination competence of mechanisms sensitive to radial, rotation and translation patterns.

Binocular disparity. Binocular disparity is known to provide a compelling sense of depth, and it has been suggested that such information could help to resolve the meaning of ambiguous retinal motion signals. In addition, cells in monkey have been found to be jointly sensitive to motion and disparity. Chapter 7 investigates whether evidence exists for such joint sensitivity in humans.
Chapter 2
General Methods

2.1 Technical Issues

The following sections address basic issues concerning stimuli design, data processing, and the control of computer equipment. Documenting the relationship between an environmental stimulus and the observer's perception is the root of psychophysics enquiry. The result of this process is frequently a psychometric function, from which parameters such as response thresholds and sensitivity are estimated. Many techniques for gathering raw data and processing it to the point of parameter estimation have been developed over the years. The first section will describe the processing methods used in this thesis to identify important trends in the data. Also detailed are the type of stimuli used here, and the reasons for their choice.

The automatic presentation of stimuli and recording of participant’s responses is commonly carried out in modern psychophysics on programmable personal computers. Recent advances in computer power, their ready availability and relatively low cost makes them an ideal choice. To use such equipment effectively to carry out experiments requires both knowledge of the capabilities and limitations of their major components, as well as the ability to control it through the development of customised software. This is the approach that has been used here, and the second section will outline some key hardware and software topics.

The final section details experimental methods that are common to many of the chapters.

2.1.1 Sampling the Observer’s Psychometric Function

It is common in psychophysics to characterise the ability of observers to perform a task over a range of stimulation, building up a more revealing picture of the competence of the mechanisms employed than would be possible with just one measurement. Several techniques exist for choosing the stimulus values to
present to the subject (e.g. the method of adjustment, the method of limits, the
method of constant stimuli, and adaptive methods. See Treutwein (1995) for a
review). In this thesis two techniques have been used: the method of constant
stimuli, and an adaptive method of stimulus selection known as ‘Quest’ (Watson
& Pelli, 1983; see also Ewen King-Smith et al., 1994). The method of constant
stimuli presents a fixed set of stimulus values to the observer that is usually
designed to sample the complete range of their performance, from guessing
through to the maximum possible. This scheme necessarily involves the
experimenter in a large degree of pilot work to ensure that the stimulus set does
indeed embrace the desired performance range. This problem is compounded by
individual differences that may dictate that some observers with different
sensitivities require a different stimulus set to others, generating more work for
the experimenter, and possibly more wasted trials for observers. Finally, this
technique has been criticised on the grounds of the relative inefficiency of its
sampling scheme (Treutwein, 1995). Sampling evenly throughout the observers’
psychometric function is held to be wasteful as some parts of the function are
more informative of its shape than others. A more efficient method would
sample more densely at those points, and less so at comparatively less
informative regions. The Quest algorithm (Watson & Pelli, 1983) offers just
such a sampling scheme, typically concentrating sampling of the stimulus space
in areas leading to high performance (e.g. 82%). This is achieved through a
Bayesian probability density function (pdf) that relates the likelihood of a
positive or negative response to stimulus values. The experimenter is only
required to supply the algorithm with an estimate of the subject’s threshold, the
algorithm then making ‘on-line’ stimulus value choices based initially on its
inbuilt pdf, and subsequently on that distribution modified by the incorporation
of the observer’s responses during the experiment.

2.1.2 Stimuli
The stimuli used throughout this thesis are referred to as Random Dot
Kinematograms (RDK). An RDK comprises a field of dots, each of which is
capable of moving with speed and directional properties that can be assigned independently of its companion's speeds and directions. A typical RDK contains dots with lifetimes limited to a short number of animation frames. Dots are 'reborn' at a new location when they expire, or when they travel outside the bounds of the (frequently circular) region of the display they are required to appear in. These properties of the RDK mean that complex patterns of motion can be created without the introduction of 'real world' geometric shapes like squares, rectangles and circles. Such shapes have properties that are not relevant to the study of low-level motion mechanisms, and may introduce undesirable cues that observers could use in place of the variables under experimental investigation. For example, chapter 4 examines how the perception of motion-in-depth varies with the spatial frequency of RDK dots making up a radially expanding optic flow. When an RDK is made up of limited lifetime dots that can be repositioned in random locations, a continuous flow field can be produced without accompanying 'object properties' intruding, or placing constraints on overall stimulus duration. If motion in depth were simulated using an approaching square for an extended period the figure would eventually fill the entire display area. Additionally its spatial frequency content would inevitably contain more power at low frequencies. Using an RDK with limited lifetime elements circumvents these problems.

Previous investigators have used different designs for the elements in their RDK's. The simplest design has a single luminance value for the body of the dot, contrasting strongly with that used for the background, whichever polarity is chosen. This design is often used for single-pixel dots where there is no opportunity to modulate the luminance profile. Graduated luminance profiles have also been used, for example Harris & Giachritsis (2000) used a triangular luminance profile in their RDK, providing a comparatively smooth transition between background and dot. The Difference of Gaussian (DoG) element design used in this thesis has been chosen for a combination of desirable properties not available with other stimuli. These features are described below.
Figure 2.1 Construction of a Gabor profile. When a sine wave (blue curve) is multiplied by a Gaussian function (red curve) the resulting windowed sine wave is referred to as a Gabor. Where the sine wave extends infinitely across space in both directions, the Gaussian window ensures that the resultant waveform is spatially restricted. The spectral composition of the Gabor is limited, being a function of the sine wave and Gaussian profile, jointly.

This discussion starts with the ‘Gabor’, essentially a sine wave whose amplitude is curtailed through multiplication by a Gaussian envelope (see figure 2.1). The properties of the Gabor stimulus have been thoroughly examined and are well known in vision science (Gabor, 1946; Marcelja, 1980; Kulikowski, Marcelja & Bishop, 1982), offering spatial localisation and spatial frequency selectivity that are believed to match the properties of frequency tuned mechanisms in primate vision (e.g. Adelson & Bergen, 1985; Field & Tolhurst, 1986). A good approximation to the Gabor is achieved through the subtraction of two Gaussian distributions, the ‘Difference of Gaussian’ (DoG) model (Rodiek, 1965; Enroth-Cugell & Robson, 1966; Enroth-Cugell et al. 1983) (see figure 2.2). As long as the Gaussian constituents differ in their standard deviations, the resulting
waveform is of similar form and spatial frequency content to the Gabor. The relationship between the standard deviations of the component waveforms determines the precise centre frequency and bandwidth of the stimulus. Figure 2.2 illustrates the result of combining Gaussian distributions of unequal deviations in one dimension. Calculation of the DoG centre frequency can be accomplished using the following formula (Clement, 1993):

\[
f = \sqrt{\frac{\ln(b \cdot s^2)}{a \left( s \cdot \sigma \cdot \pi \right)^2 \left( s^2 - 1 \right)}}
\]

where,

- \(a\) & \(b\) scale the maximum values of centre and surround Gaussians, respectively.
- \(s\) is the ratio of the standard deviations of centre and surround Gaussians.
- \(\sigma\) is the standard deviation of the surround Gaussian.
Figure 2.2 Construction of a Difference of Gaussian (DoG). When two Gaussian functions (blue and red curves) are subtracted the resulting waveform (solid green curve) is referred to as a Difference of Gaussian. The DoG is spatially restricted, like its component parts, and has a limited spectral content determined jointly by them. See text for details.

For the reductionist, then, the limited spatial frequency content of the DoG is one of its most significant advantages over other dot designs. Solid dots contain a very broad range of spatial frequencies, and are therefore incapable of targeting specific neural sub-populations, and attributing experimental results appropriately is consequently made more problematic. Of course dots are two-dimensional entities. Another advantage of using the Difference of Gaussian is its ready extension to two dimensions, which yields circular symmetry shared by other dot designs, but which is absent from 1-dimensional stimuli such as Gabors and gratings (a sine wave extended in one dimension, i.e. without the window that curtails the Gabor) that have been used extensively in low level vision research. It is also similar to the Laplacian operator (Marr, 1982; Marr &
Hildreth, 1980), in lacking orientation cues that can provide additional cues in studies where direction of motion is the attribute of interest. Another advantage of the dot over the grating is its spatial localisation. Again, this property is desirable from the point of view of matching stimulus properties to those of the mechanisms detecting the stimulus (Maffei et al., 1979; Kulikowski & Vidyasagar, 1984).

From inspection of figure 2.2 it can be seen that a significant portion of the area under the Difference of Gaussian curve is below the reference mid-line. In fact the integral of the waveform approximates zero (departure from zero is possible due to quantization errors in digital computers): the areas above and below the line are equal. This is the third important property of the DoG, it is dc balanced. If a Fourier analysis is performed on this stimulus no power exists at 0 cycles per image. This means that the mean luminance value of the dot can be matched to that of the background by choosing the mid-level of the waveform to be background luminance. Solid dots, by contrast, have a large dc component that may be detectable by low spatial-frequency mechanisms, especially where many dots are present, as in a typical RDK. In experiments that aim to require subjects to integrate over many individual dot motions the presence of a high-amplitude, widely spread luminance signal acts as a confounding variable that may be exploited unwittingly by participants (Smith, Snowden & Milne, 1994).

2.1.3 Data Processing

Behavioural research in the psychophysics tradition involves measuring how a participant’s performance varies with changes in some variable (or variables) under the control of the experimenter. The resulting psychometric function is then fitted with a model that provides a formulaic account of the relationship between performance and independent variable. From this model a single-valued measure of performance is commonly estimated. Often this statistic takes the form of a ‘threshold.’ Figure 2.3 shows a raw data set (solid diamonds) produced by a hypothetical observer detecting a low contrast visual signal in a two alternative, forced choice (2AFC) task. As the independent variable on the x-axis
increases, performance on the y-axis also rises to an asymptotic level. Identifying which of two temporal or spatial locations contains a low contrast signal, the theoretical minimum performance is 50%, and the maximum is 100%, hence the range of probability correct shown on the y-axis. The lowest possible level of performance in this example is 50% because, with two options to choose from, a guessing strategy would by chance yield the correct answer with a probability of 0.5. By chance the actual level of performance may fall above or below 50% when the stimulus is imperceptible, but tends to this value with large numbers of trials. In an experiment where an observer had to choose between four options, the theoretical range of the psychometric function would be 25% with the stimulus undetectable, to 100% when so strong as to be trivially easy to locate. The model fitted to this data set (blue curve) is sigmoidal in shape, and is in this example a cumulative Gaussian (also known as the cumulative Normal). The 'threshold' represents a theoretically interesting and representative aspect of the subject's performance. In the example of figure 2.3 the contrast (approximately 4.2%) resulting in 75% correct performance is operationally defined as the threshold, and is estimated by taking the inverse of the fitted model at the 75% level. Another attribute of the subject's performance that is sometimes quoted is the rate at which performance changes with the independent variable. This is referred to as 'sensitivity' and is captured by the slope of the model, σ.
Figure 2.3 A sigmoid function (blue curve) fitted to a hypothetical data set (blue diamonds). The graph shows how the probability of a correct answer (on the y-axis) varies with the signal strength of the independent variable on the x-axis. In this case the independent variable is the contrast of signal in a two alternative, forced choice (2AFC) detection task. Performance range: the choice between 2 options yields a minimum theoretical performance of 0.5 due to guessing, with a maximum performance of 1 at the highest stimulus values.

There are three important aspects to be considered in estimating parameters such as the threshold and slope from a subject’s raw data:

1) fitting the model,
2) assessing the 'goodness-of-fit' of the model to the data,
3) obtaining confidence limits on the parameter(s) estimated.
2.1.4 Model Fitting

As mentioned in the example, a model frequently chosen to describe observer’s performance and the one used throughout this thesis is the cumulative Gaussian. The formula for this curve is given below.

\[
f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx
\]

It is evident that there are two free parameters, \( \mu \) and \( \sigma \), that can be varied to dictate the exact shape of the waveform, although the general sigmoidal form is constant. When fitting a model of this type to data these parameters are varied to minimise some error metric that measures the overall difference between the model and the data. In the chapters that follow the error metric used is the sum of the squared differences between the empirical data and the model. The process of approximating this minimum is commonly referred to as least squares minimisation. The minimum sum of squares value can be approximated by an iterative process that adjusts the parameters \( \mu \) and \( \sigma \) in reducing amounts. The ‘solver’ routine Microsoft Excel is well suited to this task, and has been used throughout these experiments. One further refinement to the fitting process is required to reflect the use adaptive methods of selecting levels of the independent variable. Unlike the method of constant stimuli, adaptive methods such as the Quest algorithm (Watson & Pelli, 1983) used here, result in a different number of trials being collected for each level of the independent variable. This is because they progressively ‘home-in’ on certain stimulus values yielding high levels of performance that are efficient determinants of the location of the psychometric function. Consequently these stimulus values often obtain many more samples than values presented early in a block of trials. When fitting a curve to data collected in this way it is desirable to weight the fitting process more heavily by data points resulting from many samples, as these are
likely to be more accurate measures of performance. To achieve this weighting
the fitting process modulates the difference between each data point and the
model by the variability associated with the data point. For points with many
samples the variability is low, for those with few it is high. The standard
deviation for each data point is calculated individually, assuming that the
subject’s responses are drawn from the binomial distribution, and then divided
into the difference between data point and model.

Other researchers commonly use Probit analysis to fit a model to a data set.
Essentially this procedure avoids fitting a sigmoid to a sigmoidal data set by
replotting the data on log axes. This has the effect of straightening the data set,
to which a straight line can then be fitted with ease using a least squares
minimisation process of the type described above. The availability of powerful
computer hardware allows automation of the fitting procedure to take place
without going through the transformation process, hence its use here.

Conventionally fitting a model like the cumulative Gaussian to a data set
requires only two parameters (\( \mu \) and \( \sigma \), as described above) to be perturbed to
obtain the lowest error metric. However, Wichmann & Hill (2001) advocate that
one further parameter be included. The cumulative Gaussian function is
constrained to asymptote at 100%, a level of performance that an ideal observer
would achieve. However, even at very high levels of stimulus strength observers
are prone to lapses of concentration and ‘finger errors’ (i.e. making an incorrect
response despite knowing the correct answer). In such circumstances an
improved fit to the empirical data can be obtained if the asymptotic value of the
model is allowed to take on values less than 100%. In their analysis of this topic
the authors recommend that the function maximum be permitted to range
between 96% and 100%, and this strategy has been adopted throughout the
experiments reported here. One possible criticism that could be levelled at this
procedure is that any model can be made to fit a data set more closely by
increasing the number of parameters that are free to vary. Key to the decision to
add a further parameter is whether it is done merely to improve the fit, or
whether it is intended to more accurately model the observer’s behaviour.
Failure to adopt the additional parameter with a data set that incorporated lapses at high levels of performance would result in an incorrect model being applied, which would in turn lead to errors in the values estimated from it. Figure 2.4 illustrates the problem with a new example data set in which the maximum empirical performance reached was 96% correct. The blue curve is a cumulative Gaussian constrained to reach 100%, as evidenced by it overshooting the data at high stimulus values. The red curve is also a cumulative Gaussian, but has been allowed to reach asymptote at values between 96% and 100%. It can be appreciated that the distortion introduced in fitting the blue model results in an appreciably different threshold estimate from that obtained using the red curve (dashed lines estimating the stimulus value resulting in 75% correct).

Figure 2.4 Sub-asymptotic model fitting. Blue diamonds represent a new data set, similar to that of figure 2.3, except that the observer does not reach 100% performance at the highest stimulus values. Two models have been fitted to this new data set. The blue curve is constrained to reach asymptote at a probability correct of 1. The effect that this has on the 75% threshold estimate (broken lines) is evident by comparison with the red curve, whose asymptotic level was allowed to take on values between 0.96 and 1.
The blue model has been distorted, leading to a higher threshold than predicted by the red model.

2.1.5 Goodness-of-Fit

Having chosen a suitable model and fitted it to the data, it remains to be determined whether the chosen model is a plausible one for the data set recorded. Obtaining a best fit does not necessarily imply that a good model has been chosen in the first place as an 'optimal' fit can be found for any model, even one that demonstrably provides a poor description of how the observers' performance varies with the independent variable. For example, figure 2.5 shows a straight line that has been fitted to a sigmoidal data set. The error metric has been minimised, yet predictions of the subject's performance at any given level of contrast would be most likely to be inaccurate. To counter this problem a goodness-of-fit test is applied after model fitting, taking the form of a statistical prediction of the degree of mismatch between data and model. The error metric between data and model is distributed as $\chi^2$ (Press et al., 1992). Therefore by comparing the final error metric obtained during the fitting process with the $\chi^2$ distribution it can be determined with a level of confidence whether the fitted model is an adequate account of the data. The number of levels of the independent variable (9 in this example) minus the number of fitted parameters (2 parameters in this example, $\mu$ and $\sigma$) constitutes the degrees of freedom, and hence the severity of the test.
Figure 2.5 Linear fit. The data from figure 2.3 are fitted here by a straight line. Although the $\chi^2$ error metric has been minimised, the data is not fitted well by this model.
Figure 2.6 Confidence intervals. The broken blue curve was used to estimate a performance threshold at a probability correct of 0.75. The uncertainty attached to this estimate is indicated by a confidence interval (C.I.) that has a defined likelihood of containing the true threshold value. The confidence interval is obtained by fitting two additional models to the data set (green curves). The slope parameter of the green models is identical to that of the best-fitting blue model. Their position along the stimulus axis (x-axis) is determined not by minimisation of the error between the data and the model, but by aiming the optimisation procedure at a defined error value based on the error obtained in fitting the best-fit model. See text for details. The C.I. is bounded by the stimulus values found by projecting from \( p(\text{correct}) = 0.75 \) to each of the green models (broken lines).

2.1.6 Confidence Limits on the Parameters Estimated

As illustrated in figure 2.6 the data from an observer in a typical psychophysical experiment does not lie closely on the model chosen to summarise their performance. The data is said to be noisy. Therefore any response measure estimated from the model cannot be taken as an error-free estimate of the
observer's true threshold. It is necessary for the experimenter to supply a confidence interval about the estimated parameter to indicate the degree of uncertainty associated with the estimate. In the work presented in the following chapters 95% (in fact 95.4%, but referred to as 95% for convenience) confidence intervals are included, implying a 95% degree of certainty that the observers true threshold lies between the limits quoted. The method of calculating these limits is described in Press et al. (1992). Figure 2.6 shows the raw data of figure 2.3 fitted with the cumulative Gaussian model (dashed blue curve) as before. In addition two new (green) curves are shown. For the chosen threshold adopted (in this case the level of contrast resulting in 75% correct performance) each of these curves supplies one extreme of the 95% confidence interval. In the same way as the threshold is obtained, the confidence interval is computed by taking the inverse of each curve at the 75% point. This results in two contrast levels that bracket the threshold estimate, approximately 2.8% and 5.8%, as shown. The positions of these new curves are determined in the following way.

1) The slope (σ) of each new function is fixed at the final value arrived at for the best-fit model.

2) The final error metric (χ²) obtained during the fitting process of the best-fit model has its value increased by 4 (this value is specific to the case where a CI of 95.4% is required around 1 parameter, in this case μ). Call this EM4.

The mean (μ) parameter of each new curve is varied until the least sum of squares difference between the new curve and the empirical data equals EM4. By definition the best-fitting curve has a minimum χ² value. Therefore placing a new curve to either left or right of it will increase the χ² value. Doing so (in both directions) until EM4 is reached results in an interval that is likely to contain the true threshold with a probability of approximately 95%. Again this computation
has been carried out iteratively in the work reported here using the ‘solver’ algorithm in Microsoft Excel.

2.1.7 Software & Hardware
The availability of personal computers and the ability to program them has furnished vision researchers with a powerful platform on which to develop and present experiments, and to collect data from them. For the psychophysicist close control of graphics generation, display luminance and timing is required that makes exploiting computer equipment to achieve these goals a non-trivial task. Programming skills are required in an environment where most experiments are of bespoke design. A good appreciation of the limitations of computer hardware is also needed to ensure that the equipment is capable of delivering the performance asked of it. The following section will address these technical issues that are so important in conducting experiments successfully.

2.1.7.1 Software Code Generation
The generation of animations representing optic flow is central to the experiments in this thesis. To do so, software development has been carried out using the ‘C’ programming language on an Apple G4 personal computer. This language offers close control over the hardware, and is the language in which a number of software routines have been written by other psychophysics researchers. These routines execute tasks that are common to many experiments, and they can be readily incorporated into original code when a new experiment is designed. The Video Toolbox (Pelli, 1997) is a suite of such routines that have been made available to psychophysicists, and which has been used extensively here.

2.1.7.2 Hardware
The personal computer comprises three basic elements that determine the type of tasks that can be undertaken on them: the central processing unit (CPU), the graphics card and the visual display unit (VDU).
The CPU is the device that performs any calculations that are required, a process that takes a finite amount of time, depending on the number and complexity of those calculations. It may be required to generate new animations between experimental trials, involving many thousands of individual operations, in a duration that is acceptable to the observer and the experimenter. Adaptive methods of stimulus selection, such as the Quest (Watson & Pelli, 1983) algorithm used here, also perform calculations between trials, as stimulus value selection is partly dependent on the participant’s prior responses. The higher the processor speed, the more work can be undertaken in the time between trials. It is important to be aware that the CPU is not only running the user’s experimental software, but also services the Operating System (OS). The OS is a process that runs continually, occasionally making demands on CPU time to perform tasks unrelated to the experimenter’s program. In an experiment where accurate rendering and timing of animations is important, such interruptions are undesirable, and can lead to discrepancies between the intended display and the one presented. Awareness of such issues, and software control over them during programming is needed to avoid unwanted interruptions during an experiment.

The graphics card is positioned between the CPU and the visual display unit (VDU). It stores images generated by the CPU in memory chips for rapid transmission to the VDU. It is the quantity of information, and the speed with which it can be transmitted to the VDU that place one limit on the types of animation that can be displayed. To render an image on a screen involves the individual control of a matrix of picture elements, or pixels. The light emitted from each pixel on screen is determined by the quantity of electrons applied to that pixel from an electron gun, which in turn is controlled by a number stored in the graphics card. It follows that to display one image requires at least as many numbers to be stored in the graphics card as there are pixels on screen. The higher the number of pixels to be controlled, the more numbers have to be stored in the graphics card’s memory. High-resolution images are desirable, but carry
the twin penalties of using large amounts of memory, and increasing the quantity of information that has to be transmitted. A third factor, refresh rate, compounds the problems of presenting high fidelity images by reducing the amount of time available to manipulate the information to be displayed. The refresh rate is the frequency with which the display is updated with new information. High display refresh rates are desirable to simulate smooth motion since an animation is simply a series of static images displayed in rapid succession. In order to present the illusion of smooth motion the sequence of static images should be presented at a rate that exceeds the temporal resolution of the visual system. Add to this the challenge of, for example, presenting stereoscopic images and the subsequent demands on the hardware can be great. One method of presenting stereoscopic images entails presenting two slightly different versions of an image on alternate rasters of the display. When viewed through shutter glasses that are synchronised to the display, one image is seen by one eye through a transparent lens, while the other eye is obscured by an opaque lens. On the next frame the situation is reversed. In this way two different images are seen during the same temporal interval, mimicking the two slightly different views that the eyes have of a scene by virtue of their placement some 6cm apart in the head. The requirement to interleave the images means that the effective flicker rate at each eye is half that of the display. Refresh rates of 50Hz to 60Hz are widely used in television transmissions to provide acceptable fidelity, being slightly above the critical flicker frequency of most observers (Robson, 1966; Kelly, 1979). In the stereoscopic animation example a 60Hz refresh rate at each eye would demand a graphics card capable of running at 120Hz. Fortunately such cards are available, allowing stereo experiments to be performed without recourse to highly specialised and expensive equipment (e.g. Silicon Graphics).

2.1.7.3 The Visual Display Unit (VDU)
Of course the VDU to which the images are being sent must also be capable of operating at such high refresh rates, and such displays have been developed in recent years. The technology most widely used for rendering high-resolution
animations in psychophysics experiments is the Cathode Ray Tube. The 'ray' is a stream of electrons emitted from the cathode (negative electrode) of an electronic valve. The screen is itself part of this valve, and is coated with phosphor dots that fluoresce when struck by electrons, emitting light at a frequency that is a characteristic of the phosphor used. In this way colours can be produced by using several different phosphor-types emitting light at different wavelengths. The more closely packed the phosphors are the greater the resolution an image can be shown at. The light emitted from the phosphors is determined by the degree of control that can be exercised over the voltage applied to them. As noted above, the amount of light emitted from a phosphor is proportional to the voltage applied to it via the electron gun. The signal that modulates this voltage exists as a digital representation in the computer. As such it has a resolution that is dictated by the length of the binary word used for it. For example an 8-bit digital word has 256 \(2^8\) combinations, and so is able to represent a number in steps of \(\frac{1}{256}\). A 16-bit word has 65536 \(2^{16}\) combinations and can represent a number in steps of \(\frac{1}{65536}\). As there are discrete steps in the representation, neither word length is able to accurately represent intermediate numbers. However, a 16-bit word can more closely represent a given value because its steps are much smaller, reducing the quantization error. It would be advantageous to control the light emitted from the screen with as great a resolution as possible, but the word length used in most contemporary graphics cards is only 8 bits. This means that a display capable of producing 0 to 100 candelas of luminance can only do so in steps of \(\frac{100}{256}\) or 0.39 candelas. This may be insufficient for some experiments that demand finer control over luminance values, such as contrast threshold measurement. One improvement to this limitation is possible if the experimenter is content to display greyscale images, rather than colour. Pelli & Zhang (1991) proposed a method for combining the output of the three colour channels to increase the resolution of
the voltage applied to the cathode of the display. The maximum voltage to each electron gun is the same, but when passed through a potential divider resistor network can be made smaller to any degree required. This potential divider, known as the 'Video Attenuator' is used to reduce the voltage on each channel by differing amounts. This means that each channel has a different voltage step size for each change of a bit in the digital word controlling it. The voltages are then summed together to provide a single output that inherits an improved voltage resolution, equivalent to having voltages under the control of a 12-bit word. If a greyscale monitor is used with this signal then a greyscale image is the result. Connecting this output voltage to just one of the RGB (Red, Green, Blue) inputs of a colour monitor results in a monochrome image, which can be an acceptable course given the low availability and high cost of greyscale monitors. To achieve a greyscale image on a colour monitor, we have added a splitter/amplifier circuit to the Video Attenuator to multiply its single voltage output to three identical channels, which are then applied to each of the RGB inputs of the VDU.

Linearising the 'Gamma' function. For many psychophysics experiments it is essential for there to be a linear relationship between the numbers used to define an image in the computer and the luminance that the display produces. However it is not normal for a computer system to provide this desirable feature, and so the experimenter has to ensure this linear relationship is achieved through software means. This process is known as linearising the display's Gamma function. First, the existing relationship between the two is characterised by measuring the output of the display for a large range of input values. The correspondence between number and luminance can then be fitted with a function that captures the relationship between them. Finally, any non-linearity found can be compensated for by creating a lookup table through which the desired image values are passed to generate the correct luminance at the display. Age and changing environmental conditions cause drift in the specification of electronic components, so the linearisation procedure has to be repeated periodically to maintain accurate control over luminance.
2.2 Methods

The following section details methods common to many of the experimental chapters to follow. Departures from these general methods are described in the relevant chapters.

Forced choice procedures were used in all experiments, normally a single decision being made between two spatial or temporal locations: the 'two alternative, forced choice' or '2AFC' procedure. In this paradigm one location contains a signal, the other no signal, the task being to discriminate correctly between them. The locations in question could be either spatial or temporal, depending on restrictions imposed by stimulus design and the abilities of the hardware to deliver the stimuli. Each chapter details whether locations were spatial or temporal, and in the case of chapter 2 the requirement for observers to make two judgements (2*2AFC) rather than one is described more fully. Spatial or temporal locations always maintained a constant relationship to each other (e.g. spatial separation, or inter stimulus interval), but their ordering was assigned at random to prevent selective attention by the observers.

Stimuli. With one exception (see Chapter 3, Experiment 1) the stimuli used in the thesis were supra-threshold, i.e. well above luminance contrast threshold, the operationally defined minimum signal level required for a criterion level of performance. In this context the signal location referred to earlier comprises a proportion of dots sharing some organising principle, for example direction of motion, and a proportion of dots with no organisation along that same dimension. The no-signal location contains only disorganised dots, and the difficulty of distinguishing the locations is manipulated by altering the proportion of organised to disorganised dots in the signal location. This is known as the motion coherence paradigm, and is the method used in most of these experiments to establish threshold levels of performance. A motion coherence threshold is the ratio of organised dots in the signal animation to the total number of dots in that animation, and is quoted in percentage terms.
Each location comprised an annular window, curtailed at the edges by a contrast modulation. The modulation at each edge followed a cosine profile that progressively reduced the contrast at the edges from maximum to zero. This is termed a 'raised cosine' because of the extended maximum contrast central region flanked by symmetrical cosine-shaped decay.

Each animation was also contrast-modulated in time by a raised cosine window to preclude disruptive transients at the beginning and end of an animation. Animations generally lasted 1000ms, with approximately 500ms elapsing between the observers response and the subsequent trial.

A black cross was shown which subjects were required to fixate during each trial. In experiments involving two spatial locations the cross was placed halfway between locations and shown continuously. In temporal interval experiments the cross was shown between interval presentations.

Levels of the independent variables were selected using the adaptive Quest procedure (Watson & Pelli, 1983), except in Chapters 3 and 6, where the method of constant stimuli was used.

Responses. Observers initiated a block of trials using the alphanumeric keypad of the computer. If the subject perceived the left spatial location (or first temporal interval) to contain the signal, button '1' was pressed. Alternatively, button '3' would be pressed if the signal was perceived in the right spatial location (or the second temporal interval). A new trial was initiated only after their response, and there was no time pressure to answer. If the subject needed to take a break button '2' could be pressed to repeat the trial, with a new value of the independent variable being selected.

Feedback. Except where subjective judgements were required (chapters 5 and 6) auditory feedback indicating correct and incorrect judgements was given after the observers' response had been recorded. No feedback was given where subjective judgements were called for.

Viewing distance was generally 57cm or a multiple thereof, except in chapters 5 and 6.
Difference of Gaussian dots were used throughout, for the reasons detailed above, though their sizes varied between experiments. A Difference of Gaussian (DoG) dot was formed by summing two Gaussian distributions, one positive going the other negative. The standard deviation of the negative going Gaussian was 1.5 times greater than the positive. The spatial frequency bandwidth of the DoG’s was 1.77 octaves.

Dot lifetime. Dot lifetimes were tightly constrained. A dot with a lifetime of two would initially appear at a random location, jump to a new location to produce an apparent motion, then expire. A new dot would then be positioned at a random location to maintain the overall quantity of dots, and hence dot density. Synchronised birth and death of the dots was minimised by assigning each a randomly chosen elapsed lifespan at birth.

Contrast: The Michelson contrast of the dots was 0.667, and the background luminance was 50 candelas/m, except in the experiments of Chapter 7. Stimuli were viewed binocularly in a darkened laboratory.

**Equipment**

With the exception of Chapter 7 a Sony Trinitron Multiscan 400PS display was used throughout, its size being: 35(w) * 26(h) cm. A Formac GA12 graphics card was used to drive the display at a resolution of 832*624 pixels giving 24 pixels per cm. The refresh rate was typically 75Hz. The display used in Chapter 7 was a Phillips Brightview monochrome CRT (see Chapter 7 for details).

The mean luminance of the display was set to 50 cd/m², and was made a linear function of the digital control signals from the graphics card using a Minolta light meter. Pseudo 12-bit resolution was achieved using the ‘Video Attenuator’ of Pelli & Zhang (1991). The three colour guns (Red, Green and Blue, RGB) of the display were driven equally with a custom-built splitter-amplifier of our own design to produce grey level images.

Stimuli were generated on an Apple Macintosh G4 450MHz personal computer, using software routines from the Video Toolbox (Pelli, 1997). In Chapter 7 a 667MHz Apple Macintosh G4 was employed.
Chapter 3
Direction Bandwidths of Rotation, Radial and Translation Motion
Mechanisms

3.1 Abstract
The visual system is sensitive to many dimensions of the visual array, such as colour, spatial frequency and direction of motion. Neuronal mechanisms that are sensitive to a limited number of dimensions are termed 'analysers'. A useful strategy to code a wide range of values along one dimension would be to subdivide the range among several analysers within the same class. Such a scheme has been shown to exist in the processing of colour and spatial frequency, and the measurement of the tuning widths of component analysers is now well understood. Similar investigations have been carried out for mechanisms sensitive to motion direction, but this work has been largely concerned with unidirectional motion. Complex patterns of motion are prevalent in the experience of humans, and there is much evidence to suggest the existence of specialised detectors for rotation and radial patterns. It was the purpose of these experiments to measure the tuning widths of such mechanisms.
3.2 Introduction

First a note on terminology. This chapter concerns psychological mechanisms responsive to a range of directions of motion. Previous literature on this subject mixes terms in a seemingly interchangeable way, when, in fact, quite different widths are being referred to. For clarity the following terms used throughout are defined here. The ‘Bandwidth’ of a mechanism is its width at half the maximum response amplitude of the mechanism, i.e. ‘half-height.’ The ‘full-width’ of a mechanism is its width at its minimum amplitude, i.e. the most extreme extent at which the mechanism responds to a stimulus. Some authors have referred in original papers to a mechanism’s bandwidth by quoting half its width preceded by sign symbols. For example, ±30 degrees refers to a width of 60 degrees. To avoid confusion, the unsigned form will be used throughout.

Behavioural evidence has shown motion sensitive mechanisms to be restricted in the range of directions to which they are responsive (Levinson & Sekuler, 1975; Sekuler & Levinson; 1974, Sekuler, Pantle & Levinson, 1976). This range of motions is referred to as the ‘direction bandwidth’ of the mechanism and has been estimated using a variety of techniques. While researchers agree that the direction bandwidths of motion detectors are in general terms ‘broad’, seldom do two experimental paradigms yield the same quantitative estimate. For example, psychophysics experiments have proposed bandwidths as large as 150 degrees (Ball, Sekuler & Machamer, 1983), and as little as 70 degrees (Raymond, 1993). Electro-physiologists have identified that a range of directional tuning exists in V1. For example, Albright (1984) found some cells’ bandwidths to be considerably less than 40 degrees, yet many to be as large as 100 degrees and more (mean of 68 degrees), with intermediate values also being represented.

Finally, modellers of motion detection mechanisms find bandwidths of 60 degrees as a plausible minimum figure to account for human direction sensitivity (Watamanuick, Sekuler & Williams, 1989; Williams, Tweten, & Sekuler, 1991). Early experiments by Sekuler and colleagues attempted to characterize the directional sensitivity of motion sensors in terms of the independence between
mechanisms. Several techniques were used. For example, Levinson & Sekuler (1976) adapted subjects to a translating stimulus of (supra-threshold) random dots, then measured the perceived direction of motion of a subsequent test stimulus of similar design. They found that the adaptation stimulus significantly altered the perceived direction of motion of the test stimulus as long as one was within 90 degrees of the other. This finding has been interpreted to imply that independence between detectors is not achieved at less than 130 degrees of separation, (Ball, Sekuler & Machamer, 1983).

Ball & Sekuler (1979) also used supra-threshold random dot stimuli, this time in a forward masking experiment. A noise mask comprising motion in all directions was found to be capable of elevating reaction times to detect a subsequent translation test stimulus. However, when a band of directions centred on the direction of the test stimulus was removed from the noise mask, reaction times were restored. The width of the critical masking band needed to be as large as 120 degrees to fully restore reaction times.

Two studies of mutual repulsion by simultaneously presented translational motions (Marshak & Sekuler, 1979; Mather & Moulden, 1980) have also been interpreted in terms of mechanisms with broad bandwidths. The authors measured the degree to which a population of random dots could alter the perceived direction of a simultaneously presented second population as a function of the angular difference between the motions. They found that the two motion streams interacted when separated by angles up to 90 degrees. This again has been taken by Ball and colleagues to indicate that the width of directional sensors are as great as 130-150 degrees (Ball, Sekuler & Machamer, 1983).

In their own investigation of bandwidth, Ball, Sekuler & Machamer (1983) employed a double-judgement technique to estimate independence between motion-sensitive mechanisms. Subjects were required to detect in which of two animations a near-contrast threshold random dot stimulus occurred. Concurrently they had to identify whether the motion of the dots was to the left or right of vertical. The key variable was the angle of motion with respect to
(w.r.t.) vertical, which ranged from ±15 to ±90 degrees. As detection sensitivity for motion is largely unaffected by the direction of motion, it was therefore found to be invariant with the direction-of-motion variable. However, discrimination from vertical is most difficult when the direction of motion is very close to vertical, and so sensitivity for discrimination was found to improve as the angle of motion w.r.t. vertical was increased. At extreme angles of flow (>±75 degrees) the sensitivities for detection and discrimination converged, this point being taken to indicate independence between mechanisms sensing motion to the left and right of vertical. The authors found that motion signals had to be separated by as much as 150 degrees before they were stimulating independent mechanisms.

Raymond (1993) used an adaptation technique to ascertain the point of independence between neighbouring motion detectors. Motion coherence thresholds to random dot signal-in-noise stimuli were compared before and after adaptation. When adaptation and test directions were similar the thresholds were elevated, as expected. As the difference between adaptation and test directions was increased the threshold elevation diminished. Raymond found that adaptation and test directions had to differ by circa 120 degrees before thresholds remained unaffected by the adaptation stimulus. She interpreted this as reflecting a half-height bandwidth of 70-80 degrees, corresponding to a full-width of some 120 degrees (interpreted from her figure 3).

The preceding work measured the bandwidths of mechanisms sensitive to simple translational (linear) motion. However, evidence has been accumulating from psychophysics (Regan & Beverley, 1978; Morrone, Burr & Vaina, 1995; Burr, Morrone & Vaina, 1998; Snowden & Milne, 1997; Meese & Harris, 2000) and electrophysiology (Tanaka & Saito, 1989; Duffy & Wurtz, 1991; Graziano, Andersen & Snowden, 1994) supporting the existence of mechanisms responsive to more complex motions such as rotation and radial patterns. These mechanisms have also been characterized as broadly tuned in their directional sensitivity in several psychophysical studies. For example Morrone et al., (1999)
found enhanced sensitivity obtained for radial and rotation patterns compared to intermediate spirals, a result they interpreted in terms of broadly tuned cardinal detectors. Snowden & Milne (1996) similarly modeled the increased thresholds found subsequent to adaptation to complex motion patterns with Gaussian mechanisms of 40 degrees standard deviation (equivalent to 94 degrees at half-height), incorporating rotation, radial and spiral sensitive units. Meese & Harris (2001) showed that summation between individually sub-threshold stimuli declined when their directions exceeded approximately 30 degrees (when a rotating RDK pattern has each dot's nominal trajectory altered by 30 degrees a spiral pattern results, hence they can be described as being separated by 30 degrees), falling to asymptote at around 90 degrees of separation. Modelling these data, Meese & Harris (2001) proposed that mechanisms of directional tuning width 80 to 120 degrees provided an account of the results, dependent on the number of mechanisms employed. Similarly, Meese & Anderson (2002) settled on 92 degrees direction bandwidths in a continuum of rotation, spiral and radial detectors to model sub-threshold summation between those patterns of optic flow.

In summary, bandwidths of complex pattern sensors have been estimated psychophysically from differential sensitivity, either directly (Morrone et. al., 1999) or following adaptation (Snowden & Milne, 1996); or from sub-threshold summation (Meese & Harris, 2001; Meese & Anderson, 2002).

It is the aim of this paper to elaborate the understanding of the nature of rotation and radial mechanisms by estimating their directional bandwidths. Experiment 3 addresses the topic of direction bandwidths of such complex motion detectors. However, in the light of the range of estimates produced by previous authors it was felt necessary to make an independent estimate of the directional bandwidths of simple translation detectors. This would ensure that comparisons between simple and complex motion bandwidths could be made on an equivalent basis. To that end Experiments 1 & 2 address the bandwidths of
detectors responsive to simple translation motion, both local (Experiment 1) and global (Experiment 2).

Figure 3.1 Schematic of a 2*2AFC detection & discrimination trial. Stimuli were presented in two spatial locations, only one of which contained a field of 200 low contrast Difference of Gaussian dots. In the null location, right, no dots exist. The signal location, left, comprises a population of dots with identical direction components, in this case translating approximately 40 degrees to the right of vertical. Arrows shown here were not present in the dynamic stimulus and are added to assist the reader. The subjects nominated: a) in which location the signal occurred, and b) whether the signal dots flowed to the right or left of vertical.

3.3 Experiment 1. Directional Bandwidths of Local Translation Detectors
The methodology used here is closely related to that of Ball, Sekuler and Machamer (1983). Low contrast random dot stimuli were presented at angles of flow centred around vertical (see figure 3.1) and subjects were required to make both detection and discrimination judgements on each trial. As well as varying the angle of motion w.r.t. vertical, the contrast of the dots was varied in order to establish individual psychometric functions for detection and discrimination for
each angle of motion. In Ball et al’s (1983) study a near contrast threshold figure was established before hand and remained at this fixed value throughout the block of trials. Also, because several studies attest to the spatial frequency specificity of early motion mechanisms (e.g. Anderson & Burr, 1985; Watson, 1986; Ledgeway, 1996) the bandwidth of the stimuli used here was restricted through the use of filtered dots. Therefore the full-width of translation detection mechanisms was measured by establishing the angular separation between leftward and rightward motion signals required to stimulate independent mechanisms.

To appreciate the logic of the experiment refer to figure 3.2. Figure 3.2A shows the response of directional motion mechanisms to high contrast signals (in a noise-free system). However, low contrast signals depress the output of these mechanisms, as shown in 3.2B. As dots are presented at very low contrast levels it is assumed that only the most sensitive mechanism responds to the stimuli. At directions close to vertical this will be the detector whose directional selectiveness is centred on vertical (mechanism B). Precise identification of the stimulating direction is lost, according to the principle of univariance, hence discrimination performance will be poor. At directions sufficiently far from vertical to unambiguously stimulate either mechanism A or C discrimination performance will rise to a maximum. Detection performance, on the other hand, is assumed to change little across all directions and remains high at all times. The aim of the experiment was therefore to monitor how discrimination performance improved as the direction of motion w.r.t. vertical increased. When discrimination performance approaches detection performance (or reaches asymptote) points T and T’ have been identified. The full width of mechanism B is then revealed.

One problem with the foregoing analysis concerns the asymptotic nature of the Gaussian profile as it tends towards zero at its extremities. This gradual decay would appear to make estimation of a cutoff point impossible. However, if noise is taken into account, as in a real system, it is apparent that the activity of a mechanism will become unreliable as it reaches the noise threshold and a
realistic system would neglect such signals. For present purposes therefore, it can be assumed that there is a practical limit to a mechanism's width. In later calculations this width is arbitrarily assigned a value of 6 standard deviations in order to allow deductions from empirical measurements to be made.

3.2A

![Diagram showing response activity vs. direction of motion with labels A, B, C.](image-url)
Figure 3.2 Response tuning curves of directionally sensitive motion mechanisms. Figure 3.2A shows response activity to high contrast stimuli on the y-axis is plotted as a function of the stimulus direction, and is represented here as a Gaussian function. Mechanism B responds maximally to vertical motion, Mechanisms A & C respond maximally to motion slightly left and slightly right of vertical, respectively. Figure 3.2B shows the reduced response of the same mechanisms to low contrast stimuli, as used in Experiment 1. The mechanisms now do not overlap in their responses, so psychophysical performance is mediated solely by the most responsive mechanism.

3.3.1 Methods

*Stimuli*

Subjects viewed stimuli at three viewing distances, 57cm, 114cm and 230 cm. This had the effect of altering the spatial frequency of the stimuli in octave multiples. All subsequent references to measurements affected by viewing distance will be for 57cm initially, with measurements for 114cm and 230cm following in brackets.

Subjects were presented with stimuli in two spatial locations, or windows. One window contained a blank field, the other a field of low contrast isotropic
Difference of Gaussian (DoG) dots. Their centres were displaced 10.4 (5.2, 2.6) degrees either side of a central fixation cross.

Each spatial window comprised an annulus, outer diameter 10 degrees of visual angle (5, 2.5), inner diameter 0.67 degrees of visual angle (0.33, 0.17) curtailed at the edges by a symmetrical 'raised cosine' spatial window.

Dots flowed upwards at various angles w.r.t. vertical. Ten angles of motion were employed. For observer PB the angles were ± (5, 8, 12.5, 20, 31.5). For observer RW the range was ± (8, 12.5, 20, 31.5, 50) degrees, as the narrower range failed to establish a convergence of detection and discrimination performance for this observer.

Dots were programmed to jump by 0.377 degrees (0.188, 0.093) between animation frames. This figure was chosen to produce a displacement of one quarter of the wavelength of the dot’s centre frequency. The resulting speed of each dot was 3.9 degrees/s. (1.96, 0.98).

With two hundred dots programmed to occupy the spatial window dot density within the plateau of the cosine spatial window was 6.82 dots /degree² (13.64, 27.28).

For 57cm and 114cm viewing distances the contrast of the dots was varied over the range: (1.6, 2, 2.5, 3.2, 4, 5) %, using the method of constant stimuli. Greater contrasts were used in the 230cm condition, as performance was slightly poorer at this extended viewing distance: (2, 2.7, 3.8, 5.2, 7.2, 10) %.

As noted above, three conditions of spatial frequency were examined by varying the viewing distance. At 57cm viewing distance the centre frequency of the DoG's was 1.2 cycles/deg. At 114cm this figure doubled to 2.4 cycles/degree, and finally at 230cm the centre spatial frequency equated to 4.8 cycles/degree.

Procedure

A 2x2 AFC (two alternative forced choice) paradigm was employed. Subjects were required to make two 2AFC judgements on each trial. First they nominated in which of two spatial locations they perceived the motion signal. Second, they indicated whether the stimulus dots flowed to the left or to the right of vertical.
The independent variables of contrast and angle of motion were selected at random from their predefined ranges (method of constant stimuli), and were fully interleaved.

The independent variable of spatial frequency, achieved by altering the viewing distance, was explored in separate blocks of trials, their order being chosen to achieve counterbalancing.

Both subjects (the authors) were experienced psychophysics observers, having normal vision.

3.3.2 Results

Figure 3.3 shows a typical data set for observer PB viewing at 114cm, hence DoG centre frequency = 2.4 cycles/degree. The graphs reveal how the observer's performance varies with increasing contrast. Each graph shows detection and discrimination performance, the solid lines being cumulative Gaussian functions fitted using a least squares algorithm. The set of angles of motion presented was symmetrical about vertical so the data have been combined for complementary angles so that responses to, say, +5 degrees and -5 degrees are collapsed into the same data set, labelled simply 5 degrees. The succeeding panels of figure 3.3 (A to E) show that as the angle of motion is increased each discrimination curve converges on its detection counterpart. The point of convergence of the curves therefore represents half of the full width of the motion analyser centred on the vertical direction. Detection and discrimination thresholds estimated from the curves shown at figure 3.3 form a data set that is summarised in a single panel of figure 3.4. Each panel of figure 3.4 illustrates for a different spatial frequency, how such thresholds converge as the direction of motion increases departs from vertical. The detection data set was fitted with a straight line, the discrimination with a logarithmic function. The equations of these fits are shown in each panel of figure 3.4. The points of convergence were calculated by solving the equations simultaneously for a common value.
Figure 3.3 a).e) Psychometric functions showing how the probability of correct detection and discrimination performance (on the y axis) varies with contrast (on the x axis) for PB. Separate graphs are shown for each angle of motion presented w.r.t. vertical. At shallow angles of presentation discrimination performance is very poor, not reaching asymptote. As the angle of motion increases discrimination performance improves. Detection performance is good regardless of the direction of the motion.
a) PB 1.2 cycles/degree

\[ y = -0.0071 \ln(x) + 0.0449 \]

\[ y = 4 \times 10^{-5}x + 0.0147 \]

b) PB 2.4 cycles/degree

\[ y = -0.0202 \ln(x) + 0.0892 \]

\[ y = 4 \times 10^{-5}x + 0.0178 \]

c) PB 4.8 cycles/degree

\[ y = -0.0312 \ln(x) + 0.1453 \]

\[ y = 0.0001x + 0.0288 \]
\[ y = -0.0117 \ln(x) + 0.0824 \]
\[ y = -4 \times 10^{-6}x + 0.028 \]

\[ y = -0.0063 \ln(x) + 0.0467 \]
\[ y = 2 \times 10^{-6}x + 0.0174 \]

\[ y = -0.0078 \ln(x) + 0.0493 \]
\[ y = 2 \times 10^{-5}x + 0.0155 \]
Figure 3.4 a)..f) Detection and discrimination thresholds as a function of the direction of motion w.r.t. vertical. Each panel shows detection and discrimination performance for each observer [PB graphs a) to c); RW graphs d) to f)] separately under each condition of dot centre spatial frequency. Detection thresholds were fitted with a straight line, discrimination thresholds with a logarithmic function. The equations of these fits are reproduced on each figure. Points of convergence were calculated by solving the equations simultaneously for a common value, these being taken to represent the point at which responses were independent of the contribution of the vertical motion mechanism. Error bars are 95% confidence intervals.

3.3.3 Discussion

Table 3.1 summarises the results of Experiment 1. It can be seen that PB's discrimination performance did not match his detection performance until the motion stimulus presented flowed at least 30 degrees from the vertical. Thus the full-width of PB's vertical analyser is found to be circa 60 degrees (recall that the data have been collapsed about 0 degrees). The full-width of 60 degrees can now be interpreted in terms of bandwidth at half-height. Assuming the motion mechanism has a Gaussian response profile, and that 60 degrees represents 6 standard deviations of that distribution, the bandwidth is readily calculable from the formula for the cumulative Gaussian function:

Given that,

\[ f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi \sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \]

where,

- \( x \) is the independent variable vector, i.e. all levels of the I.V.
- \( \sigma \) is 1/6 * full width bandwidth, estimated empirically.
- \( \mu=0 \).

86
then the half height bandwidths can be calculated using following steps:

1) Solve for the maximum value of the function using $\mu=0$, and $\sigma$, obtained empirically.
2) Dividing this maximum value by 2 yields half the height of the Gaussian function, ‘hh’.
3) Calculate the inverse of the cumulative Gaussian function using ‘hh’ to solve for the value of the independent variable, x. Call this $X_{hh}$
4) The difference between $\mu$ and $X_{hh}$ is half the width at half height, so doubling gives the full width at half height.

<table>
<thead>
<tr>
<th>PB</th>
<th>Half width</th>
<th>Full width</th>
<th>Width at half height</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2 cycles/degree</td>
<td>50</td>
<td>100</td>
<td>39.17</td>
</tr>
<tr>
<td>2.4 cycles/degree</td>
<td>30</td>
<td>60</td>
<td>24</td>
</tr>
<tr>
<td>4.8 cycles/degree</td>
<td>35</td>
<td>70</td>
<td>27.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RW</th>
<th>Half width</th>
<th>Full width</th>
<th>Width at half height</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2 cycles/degree</td>
<td>105</td>
<td>210</td>
<td>90</td>
</tr>
<tr>
<td>2.4 cycles/degree</td>
<td>100</td>
<td>200</td>
<td>78.5</td>
</tr>
<tr>
<td>4.8 cycles/degree</td>
<td>60</td>
<td>120</td>
<td>47</td>
</tr>
</tbody>
</table>

Table 3.1 Bandwidth estimates of local translation mechanisms in degrees (column 2). Estimates of half of the full width of a directional motion analyser taken from the point of convergence of curves best-fitting detection and discrimination thresholds, see text and figure 3.4. As directions tested were symmetrical about zero degrees, the data were collapsed, removing the sign of the direction. Consequently, full width bandwidths are found by doubling the calculated widths, column 3. Bandwidths are conventionally quoted as full width at half height (column 4), which can be calculated on the assumption that the response profiles of mechanisms are Gaussian, and that their measured full width is equivalent to 6 standard deviations (see text for details).
Recall that one assumption underlying the current estimates of bandwidth is that the empirically measured width represents 6 standard deviations of a Gaussian tuning curve. The Gaussian has extended tails that, in theory at least, never reach zero. In the present context this tuning profile can be interpreted as the tendency of the motion detector to give some response to any movement within its receptive field, regardless of its direction. Avoiding this consequence of asymptotic convergence, authors such as Meese & Anderson (2002) and Meese & Harris (2001) have used cosine tuning functions in their modelling of direction bandwidths. However, electrophysiological work suggests that the Gaussian model fits directional tuning data well (Albright, 1984). Other authors have estimated bandwidth with reference to a detector's peak response (Snowden & Milne, 1996; Raymond, 1993). For example, adaptation to a direction of motion reduces the systems responsiveness to test motions of the same direction. Maximum threshold elevation occurs at a detector's peak, and the points at which thresholds are 3 dB down on the peak are taken as the half-height points. Bandwidths are therefore calculable without recourse to assumptions concerning the form of the tuning response (see Raymond 1993), although Snowden & Milne (1996) did subsequently fit a Gaussian distribution to such data. In the study by Ball, Sekuler & Machamer (1983) full tuning function widths were estimated, but no half-height deductions subsequently made. Unfortunately stopping the analysis at this point prevents comparison with bandwidths measured using -3 dB points, so it was felt worthwhile to interpret the full widths measured here in terms of half-height bandwidths. However the choice of 6 standard deviations is entirely arbitrary, and was chosen in the belief that the output of a mechanism exhibiting Gaussian tuning would have fallen to negligible levels by this stage. At very low levels of response activity noise in a biological system would render the output of the detector unreliable, suggesting a practical limit to a detector's width.
Examining Table 3.1 it can be seen that bandwidths are subject to considerable individual differences. In fact the estimates are highly dependent on the rate at which detection and discrimination curves converge (see figure 3.4), which explains the large values reported for RW. It would take only a small change in the slope of either detection or discrimination functions to make a substantial difference to the point of convergence, and hence the estimate of bandwidth. In deciding how to best characterize the way that detection and discrimination converge with increasing angle of motion another technique was considered. Hoel (1971) described a method for comparing two psychometric functions to confirm whether a single fit to both was plausible. In this way psychometric functions for detection and discrimination could be tested at each angular direction of motion to see if thresholds had converged to the same value. Unfortunately this method has two drawbacks. Firstly it ignores the trend in detection and discrimination thresholds and is dependent on only the first pair in the series to reach equality to produce a bandwidth estimate. Secondly, for some conditions and observers the convergence of the threshold curves is not complete, and so no estimate of mechanism width is possible. Therefore the point of intersection of fitted functions was chosen, allowing the trend in the thresholds to smooth out noise in the data, and to permit extrapolation in the event of a lack of convergence. As previously highlighted this technique has the consequence of producing bandwidth estimates that may be quite large if the slope of both curves is shallow, but is perhaps preferable to no estimate at all, if suitable account is taken of the vulnerability described. Nevertheless, the data shows that full width tuning functions of mechanisms responsive to translation motion may be as narrow as 60 degrees (equating to 24 degrees when expressed as a half-height bandwidth), or as great as 210 degrees (half-height=90 degrees). This range of values effectively straddles those reported previously in the literature. As detailed in the introduction, Sekuler and colleagues, and Mather & Moulden (1980) variously estimate full width tuning at >120 degrees. In comparison to Raymond’s (1993) half-height bandwidth figure of 70-80 degrees
the 24 degrees half-height bandwidth measured here still represents a decrease on even the most narrow of previously quoted figures.

Such a large discrepancy between these and previous bandwidth estimates requires consideration. In recent years a distinction has been drawn between motion mechanisms that respond to movement in a small region of visual space, and those that pool or integrate motion signals over larger areas. One influential theory postulates a hierarchy of processing, whereby many spatially restricted mechanisms feed their responses to a subsequent stage that pools their signals to produce an appreciation of the motion present globally (e.g. Maunsell & van Essen, 1983; Burr, Morrone & Vaina, 1998). If this distinction holds true the directional bandwidth of the integrative mechanism may be different from that of its contributing subordinates. Could it be that the directional bandwidth of a different class of mechanisms from those found by previous investigators has been measured? A stimulus has been developed, the motion coherence paradigm (Newsome & Pare, 1988) that ensures that observers are only able to perform judgements by pooling across many local motion signals (Downing & Movshon, 1989). In order to establish what kind of motion mechanism has been measured it was decided to repeat Experiment 1 using the motion coherence paradigm to ensure that an estimate of the directional bandwidths of global mechanisms has been obtained, with which these figures can be compared.

3.4 Experiment 2. Directional Bandwidths of Global Translation Detectors

The object of Experiment 2 was to explore the bandwidths of global motion mechanisms, and their specificity for spatial frequency. The design of Experiment 2 was similar to Experiment 1, the only changes being those required to oblige observers to integrate across many local motion signals to gain a sense of the global motion present. To ensure that the task could not be performed using one dot alone (a local judgement) the motion coherence paradigm was selected (Newsome & Pare, 1988). In this design two animations are presented, one containing noise dots, the other signal-plus-noise. The proportion of signal dots appearing in the signal-plus-noise is much less than
50%. Therefore, choosing a single dot from the signal-plus-noise animation is likely to result in the subject basing his decisions on the action of a noise dot, so producing very poor performance on the task. The dots used here were restricted in lifetime to further require the subject to integrate over many sparse local signals to obtain an appreciation of the global structure of the motion. Good performance on the task is therefore assumed to be due to a judgement applied across the entire image, i.e. a global integration over all motion signals.

3.4.1 Methods

Stimuli

As for Experiment 1, except that high contrast stimuli were employed rather than the near contrast threshold dots of Experiment 1. In the ‘signal’ animation a field of two hundred random dots actually comprised two populations. In one population all dots moved in a single direction, producing the coherent signal. In the second population each dot had a randomly selected direction component, producing incoherent motion signals. These are referred to as ‘noise’ dots. The proportion of signal dots was manipulated without altering the total number of dots in the stimulus. This independent variable was the equivalent of the contrast variable of Experiment 1, and allowed the establishment of separate psychometric functions for both detection of the signal animation, and discrimination of the signal direction. The alternative, or null location contained a population of two hundred dots where each had its direction chosen at random to produce an animation containing no coherent motion.

3.4.2 Results and Discussion

As the stimuli in this experiment were above contrast threshold interpreting the meaning of the widths measured is more problematic, and depends on how neighbouring detectors are envisaged to interact. To recap: signals will be presented symmetrically about a centre direction, seeking the point at which such signals are found to no longer interact within mechanisms sensitive to both. Referring to figure 3.5, it can be seen how such interactions are dependent on the
way in which mechanisms with overlapping response profiles are assessed, i.e. the combination rule. Let curve B represent a detector sensitive to vertical motion, whose response profile is Gaussian. Similar detectors A and C detect directions left-of-vertical and right-of-vertical, respectively. One scenario presumes that any and all mechanisms that respond to an input signal will have their outputs combined in the decision making process, i.e. the vector summation rule (Zohary, Scase & Braddock, 1996). If a signal is injected at point T1, or 'leftwards' of there, it will only excite responses from mechanism A (and nearby mechanisms to the left of A, not shown). Under these circumstances no confusions between vertical and left-of-vertical will occur, and discrimination will be maximal. The paradigm employed here will detect this point, and its partner at T1' as the extreme limits of mechanism B. Since this represents the full width of B it is apparent that half-height bandwidths are calculable, given the presumption of a Gaussian response profile (Albright, 1984; Bradley et al, 1987). It should be noted that the proximity of the neighbouring mechanisms does not influence the interpretation or calculation of bandwidth.

Figure 3.5 Neighbouring motion detectors with Gaussian response profiles. The horizontal line represents the half-height level traditionally used to characterize
mechanism bandwidths. Take curve B to be the response profile of a mechanism sensitive to vertical motion. Curve A detects motion to the left-of-vertical, and curve C detects right-of-vertical. Assuming that the empirical measure of width reflects the point at which curve B no longer contributes to direction judgements i.e. has fallen to a negligible level (scenario 1), then it can be taken to be equivalent to 6 standard deviations, from which the half-height bandwidth of curve B is easily calculated. Alternatively, under the assumption that the width measured empirically reflects the distance between the points at which curves A & C intersect with curve B (scenario 2, 'winner-take-all' combination rule), the half-height bandwidth is not calculable without knowledge of the separation between mechanisms.

An alternative scenario assumes a 'winner take all' combination rule (see Zohary, Scase & Braddick, 1996). It can be seen that signals injected at point T2, and 'leftwards' of T2, will cause only responses from detector A to be considered in the decision making process. No confusions will arise due to the contribution of B, and so discrimination performance will be maximal (equal to detection). Similarly signals at or beyond T2 'unambiguously stimulate C. The width measured empirically will therefore reflect the width of detector B at its intersection with A and C. Note, however, that as there is no knowledge of the separation of the mechanisms, the height of this intersection is unknown, and no deduction about full or half-height bandwidths can be made. However it can be claimed that the measurement would reflect the centre-to-centre distance between mechanisms, as the intersection between identical symmetrical mechanisms is directly related to their separation. Clearly the issues of width and separation are confounded. Previous authors (Watamanuick, Sekuler & Williams, 1989; Williams, Tweten, & Sekuler, 1991) modelling simple translation detectors have used the occurrence of 'metamers' to allow bandwidth estimates to be made. Metamers are perceptually indistinguishable along some dimension of interest when in fact they differ on that dimension. For example a global motion pattern made up of two directionally distinct populations of dots may look equivalent to a pattern made up of three directional components when the directions of motion are closely related. Such phenomena are taken to reveal
the processing limitations of the detectors sensitive to motion direction. Watamanuick, Sekuler & Williams (1989) proposed that a minimum of 12 mechanisms spanning the 360 degrees of possible directions was necessary to account for metameric phenomena. This assumption, plus one other supposition enabled those authors to model the directional competence of the visual system with mechanisms of 60 degrees bandwidth. Their additional assumption was that separation and bandwidth are of equal size, i.e. mechanisms intersect at their half-height points. This represents an arbitrary choice that has the advantage of allowing a minimal system architecture to be proposed, and is parsimonious. To date, information on metamers in complex motion perception is not available to inform a choice between the assumptions available. Rather than make an arbitrary decision that might lend undue weight to one of several valid architectures, the strongest conclusion to draw is that the bandwidth of mechanism B is as broad, or possibly broader than the empirical measures that can be made using this technique. As high degrees of accuracy can be obtained with a large number of mechanisms that have very narrow widths, or fewer broadly tuned mechanisms (e.g. Snowden, Treue & Andersen, 1992) there is no a priori reason to prefer either narrow or broad tuning (see also Pouget, Deneve & Ducom, 1999).

To summarise, a winner take all combination rule seems less likely from an intuitive point of view, since it involves neglecting information that has been gathered. More pragmatically, it is difficult to interpret the empirical widths in terms of half-height bandwidths, without further recourse to strong assumptions concerning mechanism separation or number. For these reasons it has been assumed that simultaneously responsive mechanisms have their outputs pooled in some fashion. This allows the interpretation of these results as reflecting full width mechanism bandwidths, from which half-height bandwidths may be calculated with only one additional weak assumption: that mechanisms have a Gaussian response profile.
Figure 3.6 Schematic of a 2*2AFC detection & discrimination trial. Stimuli were presented in two spatial locations, both of which contained a field of 200 Difference of Gaussian dots. In the 'noise' location, right, all dots have a direction component chosen at random to produce incoherent motion signals. The signal location (signal + noise), left, comprises two populations of dots. One population of dots shares an identical direction component, in this case translating approximately 40 degrees to the right of vertical (filled dots). The remaining dots each have a randomly chosen direction component (unfilled). Arrows and filling shown here were not present in the dynamic stimulus and are added to assist the reader. The subjects nominated: a) in which location the signal occurred, and b) whether the signal dots flowed to the right or left of vertical.

All other methodological details were the same as in Experiment 1. In the motion coherence paradigm the ratio of signal dots is manipulated to establish thresholds for task performance. The following set of signal ratios was presented: (10, 16, 25, 40, 63, 100)%.

The process used to determine estimates for half-height bandwidths in Experiment 1 was applied to the results of Experiment 2. Table 3.2 shows for
both observers how these bandwidth estimates vary with the centre spatial frequency of the stimulus dots.

<table>
<thead>
<tr>
<th>PB</th>
<th>Half width (degrees)</th>
<th>Full width (degrees)</th>
<th>Width at half height (degrees)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2 cycles/degree</td>
<td>25</td>
<td>50</td>
<td>19.6</td>
</tr>
<tr>
<td>2.4 cycles/degree</td>
<td>25</td>
<td>50</td>
<td>19.6</td>
</tr>
<tr>
<td>4.8 cycles/degree</td>
<td>30</td>
<td>60</td>
<td>23.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RW</th>
<th>Half width (degrees)</th>
<th>Full width (degrees)</th>
<th>Width at half height (degrees)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2 cycles/degree</td>
<td>20</td>
<td>40</td>
<td>15.7</td>
</tr>
<tr>
<td>2.4 cycles/degree</td>
<td>15</td>
<td>30</td>
<td>11.75</td>
</tr>
<tr>
<td>4.8 cycles/degree</td>
<td>15</td>
<td>30</td>
<td>11.75</td>
</tr>
</tbody>
</table>

Table 3.2 Bandwidth estimates of global translation mechanisms in degrees. Estimates of half of the full width of a directional motion analyser taken from the point of convergence of best-fitting detection and discrimination curves. As directions tested were symmetrical about zero degrees, the data were collapsed, removing the sign of the direction. Consequently, full width bandwidths (column 2) are found by doubling the calculated widths. Bandwidths are conventionally quoted as full width at half-height, which can be calculated on the assumption that the response profiles of mechanisms are Gaussian, and that their measured full width is equivalent to 6 standard deviations (see text for details).

It is apparent from Table 3.2 that there is no consistent indication that direction bandwidths of global motion mechanisms are dependent on spatial frequency. However, comparing Table 3.2 with Table 3.1 it is clear that the breadth of tuning measured with the present stimuli is lower than in Experiment 1. These results differ sufficiently to suggest that the previous figures are likely to reflect the bandwidths of local mechanisms, as intended.

This still leaves the problem of explaining why these global mechanism bandwidths are so much narrower than indicated by previous investigators'
work. In her paper on this subject Raymond (1993) suggested that much work prior to that date had relied on supra-threshold stimuli, and that this may have led to excessive stimulation of neighbouring mechanisms, and consequently their inadvertent inclusion in the bandwidth estimate. Only Ball, Sekuler & Machamer (1983) used low contrast stimuli, on the assumption that they stimulated only the most sensitive mechanism, although those authors set contrast near to a threshold that was established prior to the experimental trials, and remained at this level throughout. Raymond’s criticism implies that their chosen level of contrast may have been too high to ensure that they stimulated only the mechanism intended. However establishing thresholds independently for each direction of motion has led to similarly broad bandwidths in Experiment 1, but narrow bandwidths in Experiment 2. An alternative explanation is that the difference in bandwidths reflects the division between local and global mechanisms.

Raymond (1993) used high contrast motion coherence stimuli, similar to those used here. Motion coherence thresholds were measured subsequent to adaptation by a motion stimulus. Subjects’ perceptual experience of global motion was used in a ‘yes-no’ task to establish the range of directions over which thresholds were elevated compared to a ‘no adaptation’ condition. The use of an adaptation paradigm in combination with a yes-no task may help explain why the present data differ from those of Raymond (1993). The adaptation technique is not renowned for its precision, and recent work (e.g. Foley & Chen, 1997) highlights that the interpretation of its effects is less than straightforward, especially in terms of the ‘fatigue’ assumption. Also, use of the yes-no paradigm is likely to lead to more conservative reporting of the presence of global motion, probably leading to more generous and further reaching adaptation effects than the forced choice procedure employed here, and consequently larger bandwidth estimates.

For now it is enough to recognize that the tuning of global motion mechanisms appears to be somewhat narrower than even Raymond (1993) found.

As previously pointed out, the nature of the mechanisms measured (i.e. local versus global) is an important factor in interpreting quoted bandwidths. For
some stimuli the current theory of hierarchical motion processing informs which mechanisms are likely to be relied upon, whereas for others the situation is undetermined. For example, the motion coherence paradigm used here and by Raymond (1993) is widely thought to stress the contribution of global mechanisms, whereas very low contrast stimuli, or those occupying only a small region of visual space are presumed to target local motion mechanisms (Morrone et al., 1995) (e.g. Experiment 1 and Ball, Sekuler & Machamer, 1983). Table 3.3 summarises the literature on bandwidth estimates according to the level of processing presumed to mediate performance. In order to facilitate comparisons between studies all estimates have been quoted as bandwidths at half-height, employing the same assumptions and technique used in Experiment 1 to convert from the full widths often quoted. Unfortunately, a clear pattern does not emerge from the data. Prior to Raymond (1993) a consensus of over 50 degrees is consistent with the results found here for local mechanisms. However Raymond’s 47 degrees bandwidths are similar to those, despite the use of the motion coherence paradigm targeting global mechanisms. Yet within the study reported here there is a clear drop in bandwidth with the move to the motion coherence stimulus. It is worth noting that Raymond did claim narrower bandwidths than previous studies, although no allowance was made for the difference between full and half height estimates.
<table>
<thead>
<tr>
<th>Study</th>
<th>Full width bandwidth estimate (degrees)</th>
<th>Half-height bandwidth estimate (degrees)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levinson &amp; Sekuler 1976</td>
<td>130</td>
<td>52*</td>
</tr>
<tr>
<td>Ball &amp; Sekuler 1979</td>
<td>135</td>
<td>53*</td>
</tr>
<tr>
<td>Marshak &amp; Sekuler 1979</td>
<td>135</td>
<td>53*</td>
</tr>
<tr>
<td>Mather &amp; Moulden 1980</td>
<td>130</td>
<td>52*</td>
</tr>
<tr>
<td>Ball, Sekuler &amp; Machamer 1983</td>
<td>120</td>
<td>47*</td>
</tr>
<tr>
<td>Raymond 1993</td>
<td>120</td>
<td>47*</td>
</tr>
<tr>
<td>Watson &amp; Bex</td>
<td>60-210</td>
<td>24-90</td>
</tr>
<tr>
<td>Watson &amp; Bex</td>
<td>30-60</td>
<td>12-24</td>
</tr>
</tbody>
</table>

Table 3.3 Summary of the findings of studies where the breadth of the tuning functions of motion sensitive mechanisms have been estimated or implied.

* calculated from the authors’ width estimates, which has been assumed to reflect the full width of a single mechanism, whose response profile is presumed to be Gaussian.

Stimulus types and procedures in the experiments at table 3.3 were as follows:

Levinson & Sekuler (1976). Two wide-fields of high-contrast, unidirectional random dots were presented. Observers adapted to the first stimulus (speed=4 degrees/s) for 3 minutes, then nominated the perceived direction of a subsequent field of dots.

Ball & Sekuler (1979). An omni-directional wide-field masking stimulus (filtered of some directions) of high contrast random dots changed unpredictably to a wide-field of unidirectional dots. Observer’s reaction times to the change were recorded.
Marshak & Sekuler (1979). Two divergent streams of high contrast dots were shown moving at 4 degrees/s for 1 second. After presentation observers reported the perceived angle between the streams.

Mather & Moulden (1980). Two divergent wide-field populations of dynamic noise (‘test’ field and ‘inducing’ field) moving at 2.34 and 2.58 degrees/s were superimposed. Observers reported the perceived direction of the ‘test’ population.

Ball, Sekuler & Machamer (1983). A wide field of dots at/near contrast threshold were shown in one of two temporal intervals, streaming at angles to the left or right of vertical. Observers detected the interval, and discriminated the direction of flow, left or right of vertical.

Raymond (1993). A wide-field of either unidirectional or omni-directional high contrast random dots (speed=0.92 degrees/s) was presented after adaptation to uni-directional motion. Observers discriminated which type of motion was presented post adaptation. Motion coherence thresholds were recorded as a function of the difference between adaptation and test directions.

Watson & Bex. As described in this chapter.

3.5 Experiment 3. Directional Bandwidths of Rotation and Expansion Detectors

Experiment 3 was designed to probe the directional bandwidths of more complex forms of motion. Whereas the motion signal in Experiments 1 and 2 was a simple translation involving only one direction of motion, Experiment 3 investigated the bandwidths of mechanisms sensitive to the multiple directions of rotational and radial stimuli.

Much evidence suggests the presence of detectors specialised for these patterns in the human visual system. Morrone, Burr & Vaina (1995) found that sensitivity to rotational and radial motion increased in a square root relationship with stimulus area when non-signal sectors of their display were left blank. With those same sectors filled by noise dots, sensitivity increased linearly with stimulus area. They argue that these properties are consistent with an ideal
integrator model with specialised detectors tuned to radial and rotational motion. Morrone et. al. (1999) found that sensitivity to complex patterns ranging from rotation, through spiral to radial motion was greatest to rotation and radial motion, suggesting detectors tuned to these cardinal directions. If correct this would imply that cardinal detectors would need to be broadly tuned to mediate spiral motion perception. However the study did not address the bandwidths of cardinal detectors directly. Snowden & Milne (1997) provided evidence for specialised complex motion mechanisms, sensitive to rotation and radial patterns. For example, their rotation stimulus comprised two sectors of a virtual window, leaving two interleaved sectors un-stimulated. After a period of adaptation a test stimulus was subsequently positioned in the un-stimulated areas resulting in a 'phantom motion after-effect' being perceived by subjects. Similar effects were found for radial stimuli. Since this phenomenon cannot be explained by the adaptation of local motion detectors (different populations of local detectors were stimulated at adaptation and at test suggesting that adaptation effects could not be mediated by these independent sensors) the authors conclude that specific wide-field rotation detectors 'generalize' the adaptation after-effects throughout their entire receptive field, even to areas not initially stimulated.

Speaking directly to the issue of the bandwidths of radial and rotation detectors, Meese & Harris (2000) conducted a series of sub-threshold summation experiments. Two populations of sub-threshold RDK motion signals were presented simultaneously. For example a pure rotational motion was shown in the presence of a spiral motion to see how the pair might sum to produce a supra-threshold signal. The authors documented how sensitivity to the collective signal varied with the angular difference between the component signals. It was found that the signals interacted at large angular differences, indicating that detection mechanisms are broadly tuned. Meese & Harris (2000) modelled their data and suggest plausible half-height bandwidths of between 80 and 120 degrees, depending on detector numbers and separation. Meese & Anderson (2002) used a similar technique, proposing bandwidths of circa 92 degrees.
To extend the work of Meese and Harris (2000) into the supra-threshold domain a motion coherence task was again chosen to establish detection and discrimination functions for a range of conditions. From these functions the bandwidths of putative rotation and radiation detectors could be estimated in a similar way to the previous experiments.

The concept of bandwidth in a multiple direction stimulus can be understood in the following way. Consider a mechanism sensitive to rotational motion. Refer to figure 3.7 A. If dots followed the trajectory indicated by the blue arrows a pure rotational motion would be generated. The blue arrows represent the ideal trajectory for rotation, given their spatial location, and each can be considered a reference angle of zero for the following discussion. Any departure from these ideal trajectories can be quantified in angular terms, as shown by the red arrows in the figure. The resulting pattern would be an inward spiral, which could be described as having an angular offset of, say, -30 degrees. Now consider figure 3.8 which shows response profiles of a mechanism sensitive to rotation. The x-axis is calibrated in angular offset from ideal rotation. The response profile of our rotation sensitive mechanism is therefore centred on zero degrees, and its ideal stimulus is represented by arrow A at zero degrees. Arrow B represents the spiral stimulus of figure 3.7 A i.e. -30 degrees. Clearly the mechanism’s response profile indicates that it would respond to this spiral, and its bandwidth, as defined on this dimension, is the range of spiral stimuli to which it would respond.

To measure the range of (spiral) directions around pure rotation that a rotation mechanism is sensitive to, a variable angular component was added to the nominal rotation vector. Depending on the sign of that component this would have the effect of producing an expanding or contracting spiral (see figure 3.7A). The detection judgement involved nominating in which of two spatial locations a coherent motion signal occurred. For the discrimination judgement the observer had to nominate whether the spiral signal was perceived to expand
or contract. Both clockwise and counter-clockwise rotations were used, being chosen at random for each animation and each trial.

Similarly, to measure the range of (spiral) directions around pure radial motion that a radial motion mechanism is sensitive to, a variable angular component was added to a nominal expansion or contraction pattern (figure 3.10 refers).

Figure 3.7 Schematic of a detection and discrimination trial. Both spatial locations contain a 200 dot RDK. Panel B has random motion as indicated. In Panel A if the dots followed the blue arrows they would form a pure rotation. In fact an angular offset is added to this ideal rotation trajectory such that dots follow the red arrows, simulating an inward spiral motion. Several levels of the angular offset variable were explored, a motion coherence threshold for both detection and discrimination being established for each. Subjects were required to detect in which location the spiral appeared (detection), and whether it was an inward or outward spiral (discrimination).
Figure 3.8. The response function of a hypothetical detector tuned to rotation motion, modelled as a Gaussian function. The blue arrow indicates the stimulation presented by a set of ideal local motions forming a rotating pattern (see text and figure 3.7). The red arrow represents the stimulation caused by a spiral pattern, formed by perturbing the ideal rotation trajectories by an angular offset of -30 degrees. The bandwidth of this detector allows it to respond to spirals providing that they are formed by local motion vectors within a limited number of degrees of the ideal rotation vectors.

3.5.1 Methods

Stimuli

In the description that follows rotation stimuli will be referred to for simplicity, but the same details apply to radial motion with appropriate substitution of terms. Rotation and radial bandwidths were explored in different blocks of trials. As in Experiment 2, high contrast signal-in-noise stimuli were employed rather than the near contrast threshold dots of Experiment 1. Two spatial locations were presented in the 2*2AFC task, as before. In the ‘signal’ location a field of two hundred random dots (difference of Gaussian) comprised two populations. In one population all dots were initially assigned a direction consistent with a rigid rotation about the centre of the stimulus window, producing the coherent signal. Before being displayed these dots had their trajectories perturbed (see below),
however the speeds calculated for rigid rotation were retained, such that a
smooth speed gradient was produced. Therefore increasing eccentricity is
correlated with increasing speed (see figure 3.9). In the second population each
dot had a randomly selected direction component, producing incoherent motion
signals. These are referred to as ‘noise’ dots. The ratio of signal dots was
manipulated without altering the total number of dots in the stimulus by
reducing the number of noise dots in tandem. This independent variable was the
equivalent of the contrast variable of Experiment 1, and allowed separate
psychometric functions for both detection and discrimination to be established.
The alternative location contained a population of two hundred dots where each
had its direction chosen at random to produce a ‘noise’ animation containing no
coherent motion. Their speeds varied with eccentricity to match those of the dots
simulating rigid rotation.

Figure 3.9 Rigid rotation. In the diagram a two frame apparent motion stimulus is
depicted, simulating a rigid rotation. All three dots must travel from their start point to
their destination in the same time. As time is constant the length of the arrows is
proportional to the speed of each dot. It is apparent that the speed of each dot in such a
pattern is related to its eccentricity from the centre of rotation.
In Experiments 1 and 2 signal dots translated near vertical, the precise angle w.r.t. vertical being an independent variable. In this experiment the signal dots nominally formed a rotational motion (as described above). The independent variable in this case was an angular offset added to the rotation vector. This resulted in a spiral motion, with levels of this I.V. selected on each trial randomly from the set: ± (6, 10, 16.5, 27, 45.5) degrees.

The excursions of spiralling and noise dots in the signal animation were scaled to their (initial) eccentricity producing speed gradients appropriate for rigid rotation. The speeds of noise dots in the alternative animation were similarly scaled with eccentricity to ensure that the same distribution of speeds was present in both locations.

Since Experiments 1 and 2 indicated that spatial frequency did not have a systematic effect in determining directional bandwidths, and existing evidence suggests that spatial frequency tuning of global mechanisms is broad (Yang & Blake, 1994; Bex & Dakin, 2002) only dots with centre frequency 2.4 cycles per degree were used in this experiment. This was achieved by fixing the viewing distance at 114cm.
Figure 3.10 Schematic of a detection and discrimination trial. Both spatial locations contain a 200 dot RDK. Panel B has random motion as indicated. In Panel A if the dots followed the blue arrows they would form a pure expansion. In fact an angular offset is added to this ideal rotation trajectory such that dots follow the red arrows, simulating clockwise spiral motion. Several levels of the angular offset variable were explored, a motion coherence threshold for both detection and discrimination being established for each. Subjects were required to detect in which location the spiral appeared (detection), and whether it was clockwise or anti-clockwise (discrimination).

Procedure
As for Experiments 1 and 2, except that in the rotation experiment subjects had to detect in which location a spiral appeared, and to discriminate whether a contracting or expanding spiral had been present. In the radial experiment the location containing a spiral was detected, and the direction of spiral, clockwise or counter-clockwise was discriminated.

3.5.2 Results
The process used to determine estimates for half-height bandwidths in Experiments 1 and 2 was applied to the results of Experiment 3. Table 3.4 shows the bandwidths obtained for both observers.
### Table 3.4 Bandwidth estimates of global radial and rotation mechanisms in degrees.

Estimates of half of the full width of a directional motion analyser taken from the point of convergence of best-fitting detection and discrimination curves, see Experiment 1 for details. As directions tested were symmetrical about the nominal motion type (rotation or radial motion), the data were collapsed, removing the sign of the direction. Consequently, full width bandwidths are found by doubling the calculated widths, column 2. Bandwidths are conventionally quoted as full width at half height, which can be calculated on the assumption that the response profiles of mechanisms are Gaussian, and that their measured full width is equivalent to 6 standard deviations (see Experiment 1 for details).

<table>
<thead>
<tr>
<th>PB</th>
<th>Half width (degrees)</th>
<th>Full width (degrees)</th>
<th>Width at half height (degrees)</th>
</tr>
</thead>
<tbody>
<tr>
<td>radial</td>
<td>25</td>
<td>50</td>
<td>19.6</td>
</tr>
<tr>
<td>rotation</td>
<td>25</td>
<td>50</td>
<td>19.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RW</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>radial</td>
<td>20</td>
<td>40</td>
<td>15.7</td>
</tr>
<tr>
<td>rotation</td>
<td>35</td>
<td>70</td>
<td>27.6</td>
</tr>
</tbody>
</table>

#### 3.6 General Discussion

Mechanisms whose ability to detect motion is spatially restricted (i.e. 'local') were shown in Experiment 1 to have bandwidths somewhat narrower than has previously been estimated. Experiment 2 sought to estimate bandwidths of mechanisms integrating over larger regions of space (i.e. 'globally') to confirm that the findings of Experiment 1 did not inadvertently reflect the action of global mechanisms. Why doubt that the stimuli in the first experiment excited only local mechanisms? The first and most obvious observation is that those stimuli (see figure 3.1) extended over large regions of space, and so would certainly be candidates to be integrated into a global percept. In answer to this criticism it is sometimes assumed that this global integration does not take place at the very low levels of contrast used. To detect the presence of a stimulus at
contrast threshold is presumed to require, minimally, only one of the many dots available. Still, there are reasons to question this interpretation. Firstly, there is no way of confirming or ensuring that one dot is being used to mediate the task. Secondly, the point of contrast threshold is an interpolated point between seeing all of the stimulus, and seeing none of it. It may not be possible to interpret it as equivalent to seeing a single dot. For these reasons the possibility that stimuli at low contrast levels excite global motion mechanisms must be entertained. Perhaps the only way to avoid integration over many local motions is to use only one dot in the stimulus. Nevertheless, since the latter experiment yielded substantially narrower bandwidths than the former, two different levels of processing appear to have been tapped.

Accepting that global mechanism bandwidths may be as narrow as 16 degrees (Experiment 3), these results bear on the ‘basis set’ theory of complex motion processing. Some authors (e.g. Morrone et al, 1999; Burr, Badcock and Ross, 2001) have cited evidence for ‘cardinal detectors’, that is those tuned for the orthogonal directional patterns of rotation and radial motion. If these and only these detectors exist it would be possible to code for intermediate spiral motions via a combination of this basis set, provided that basis components were broad in their tuning. This concept is therefore incompatible with these results as such narrowly tuned mechanisms would not provide sufficient coverage of intermediate directions to be simultaneously stimulated by a spiral falling midway between. Consequently this work represents further support for those favouring the existence of sensors coding intermediate directions directly (e.g. Snowden & Milne, 1996; Meese & Anderson, 2002), a position supported by the known physiology (e.g. Duffy & Wurtz, 1991; Graziano, Andersen & Snowden, 1994), and reflected in the emergent properties of MST simulation (Beardsley & Vaina, 1998).

It has to be acknowledged that the model used to calculate bandwidths is critical to the values suggested. While the framework employed has the benefit of
simplicity, other, more realistic models exist. For example Beardsley & Vaina (2001) recently simulated a biologically inspired architecture that could be responsible for mediating discrimination performance between perturbed optic flow patterns similar to those presented here. The key objective of their study was to investigate how lateral interconnections within an MST-like layer might be involved in achieving the discrimination performance of human observers. They found that lateral inhibition between mechanisms tuned to a range of optic flow patterns (radial through spiral to rotation) was required to obtain the most accurate simulations. Such interactions are absent from the conceptual framework used to interpret these data, and their inclusion might change the estimated bandwidths significantly.

Viewing distance manipulation.

In the foregoing experiments the spatial frequency of the DoG elements was manipulated by altering the viewing distance. In addition to changing the spatial frequency this strategy had the effect of varying both the size of the stimulus, and the retinal speed of the DoG elements. As these changes were confounded with the spatial frequency manipulation it is not possible to say whether they exaggerated or mitigated the inconsistent bandwidth figures recorded in these experiments. Changes in size have been found to affect perceived speed (Snowden, 1999), however it is not known whether changes in either of these attributes affects the ability to discriminate the direction of complex motion. Field size has been shown to influence discrimination between opposite directions of unidirectional motion (Burr, Morrone & Vaina, 1998). It is possible that too small a stimulus, i.e. one that is not matched to the receptive field size of the detecting neurones would lead to less accurate estimates of direction, introducing noise to the data.

Based on the ability of observers to discriminate between opposing unidirectional global motion, Edwards, Badcock & Smith (1998) proposed that two global motion systems existed, each differently tuned for speed (see also De Bruyn & Orban, 1988). If this finding were to generalise to complex optic flow
patterns then it is conceivable that altering the retinal size and hence speed of our patterns could have stimulated different systems. However the consequences of this for direction discrimination in complex motion patterns are as yet unknown and remain to be explored.
Chapter 4
The Dependency of Sensitivity on Spatial Frequency in Radial Flow

4.1 Abstract
Relative motion between an observer and an approaching textured object produces a radially expanding pattern of optic flow on the retina. Not only does the retinal projection of the object grow in size, its local features grow and diverge from each other over time, providing cues to the presence of motion-in-depth. Previous studies examining the relative importance of size change and divergence rate have concluded that size change plays only a minor role in 3d motion perception compared to divergence. However recent evidence (Schraeter, Knill & Simoncelli, 2000) suggests that spatial scale has an important role to play in the perception of motion-in-depth. To explore further the influence of spatial frequency on the perception of 3d motion two experiments were performed. Motion coherence thresholds to detect radial expansion in a random dot stimulus were measured across a range of speeds. The rate at which individual filtered dots grew was manipulated while dot divergence remained constant.
4.2 Introduction

When an observer watches an approaching object, such as a football, the image projected on the retina grows larger, and any markings on the ball will, in retinal terms, become more widely separated and will simultaneously increase in size. All of these cues may be useful in interpreting the retinal signals as being produced by an object moving in depth. Previous research has confirmed the effectiveness of these attributes in stimulating a sense of motion-in-depth.

Object size change: Regan & Hamstra (1993) established that rate of change of size could be used effectively by subjects in discriminating the time to contact (TTC) of a uniform square used to simulate an object approaching along the line of sight. This ability to discriminate TTC according to rate of size change was also found to hold for a peripherally presented square by Regan & Vincent (1995) using similar stimuli and methodology. Todd (1981) presented two expanding squares, each defined by individual dots (the dots remained of fixed size as the squares grew larger). Subjects were required to discriminate which square would reach them first, a task that they were readily able to do with high accuracy, suggesting that they could use rate of size change to discriminate the relative TTC of the simulated objects.

Object size and texture: Beverley & Regan (1983) presented a textured square stimulus. The texture could expand or contract, while the overall size of the square could change in size congruently, change in size in the opposite direction, or could remain fixed. Their method was to adapt subjects to these various combinations of object and texture size changes, and measure the amount of motion required to null the motion after-effect (MAE) induced. The test pattern was an un-textured square whose direction of size change matched that of the adapting stimulus, and therefore was opposite to the after-effect. They found that after-effects were maximal when texture and square size changed in tandem; that texture size change alone (i.e. no overall object size change) was less effective at inducing a motion after-effect, and that conflicting size and texture changes
could result in abolition of the motion after-effect. This clearly demonstrates the interaction of both sources of information on a metric (the MAE) that is widely assumed to indicate the selective response of the motion system. Vincent & Regan (1997) documented the performance of subjects estimating TTC when presented with a simulation of an approaching textured square whose size and texture growth rates were manipulated independently. They found that TTC discrimination thresholds were little affected by mismatching the two attributes, but that errors in estimating TTC were low when square growth and texture growth were congruent, but increased with the degree of mismatch in their growth rates.

Size, separation and texture: Gray & Regan (1999) investigated the effect of dot size on the accuracy of TTC estimates. While object size and texture divergence increased in tandem, dots could either grow veridically, or remain of fixed size. Their stimulus was a square figure made up of evenly spaced dots with a Gaussian luminance profile. Estimates of TTC were found to be most accurate when all three attributes were made congruent. When texture elements (circular dots) were small, fixing their size had little effect on errors in TTC estimation, but this was not true of large dots, which produced large over-estimations in TTC.

Texture and separation: By presenting a random dot kinematogram (RDK) within a constant diameter window Harris & Giachritsis (2000) were able to remove the overall stimulus size as a cue to motion-in-depth, leaving only dot (or texture element) growth and dot divergence rates as relevant variables. Again TTC was used to assess the relative contribution of dot growth and dot divergence. Except when growth and divergence signalled opposite directions of motion-in-depth, they found that performance was little affected when element growth rate was manipulated, and concluded that the size change of their RDK dots had only a small part to play in determining the perception of retinal flow compared to the relative motion among dots.
The consensus from the foregoing studies is that element size change has a limited influence on the perception of stimuli simulating looming, with overall size change and element divergence rates remaining as the most significant cues to motion-in-depth. Hence local directional signals seem more prominent than size change in determining quantitative performance in motion-in-depth tasks.

Current models of local motion detection postulate spatially and directionally restricted mechanisms, capable only of sensing motion within a small region of visual space and within a limited range of directions (e.g. Albright, 1984). In order to detect wide field optic flow created by events such as forward motion it has been suggested that these local motion signals are integrated in a second stage of processing that allows interaction between local signals. In this way global motion detectors are constructed from simple elements (e.g. Perrone & Stone 1998). However, a further feature of models of local detectors is their limited spatial frequency bandwidth (Adelson & Bergen, 1985; Harris, 1986; Van Santen & Sperling, 1984; Watson & Ahumada, 1985). Given this specificity of response at the local level, it follows from the hierarchy described above that global motion mechanisms might be restricted by the spatial frequency specificity of their local precursors. Support for this idea was provided by Ledgeway (1996) who presented wide field translation in two frame RDK's. Observers were asked to discriminate the global direction of motion. Each frame of the animation was convolved with its own filter, discrimination only being possible when filters were separated by less than an octave in frequency space. This result was taken to show that local mechanisms restricted in their frequency tuning constrained the global perception of the direction of motion.

Using random dots, Yang & Blake (1994) provided evidence that the spatial frequency tuning of global mechanisms is quite broad, measuring a bandwidth of around 2.4 octaves at half-height. In their detection task, where a translation signal immersed in dynamic noise was discriminated from random motion in dynamic noise, dynamic noise dots were convolved with a different filter to that used on all other dots. Yet even when filter centre frequencies were widely
spaced discrimination performance was reduced by the noise-mask, implying that a wide range of spatial frequencies was being integrated in global translation detection.

The preceding tasks used translating stimuli to investigate the consequence of using narrow band stimuli on the perception of global motion. The effect on complex forms of global motion, such as radial flow has not received much attention. To test whether spatial scale plays a part in the detection of radial motion two experiments were performed. In both experiments the spatial structure of elements making up a random dot field simulating radial optic flow was manipulated. In the first experiment a radially expanding or contracting RDK was presented whose elements were circularly symmetric difference of Gaussian dots (DoG's). The spatial frequency bandwidth of these dots is limited, and changes in tandem with their retinal size, therefore altering element growth rate affects spatial scale change. Four rates of element growth were presented: the zero growth condition involved no size or spatial frequency change between successive samples of a dot's trajectory. The veridical growth condition saw the dots grow at a rate commensurate with the rate of dot divergence (commensurate in the case of the approach of a fronto-parallel plane). The third level of growth was double the veridical rate, and the last was four times the veridical rate. The speed of the optic flow pattern was varied to see if any effect of spatial scale differed with this parameter. Three conditions of velocity were explored.

In the first experiment the spatial scale of the optic flow pattern was manipulated by changing the size of a spatially limited dot, the difference of Gaussian. The spatial scale manipulation was dependent on the association between size and spatial frequency content. In a second experiment a stimulus was devised where overall size and spatial frequency content were dissociated. In this way it was possible to explore the effect of stimulus size on the detection of optic flow, while maintaining spatial scale constant.
Figure 4.1 Schematic of a 2AFC detection trial. Stimuli were presented in two spatial locations, both of which contained a field of 75 Difference of Gaussian dots. In the 'noise' location, right, all dots have a direction component chosen at random to produce incoherent motion signals. The signal location (signal + noise), left, comprises two populations of dots. One population of dots referred to as 'signal' (filled) are arranged to produce a radially structured pattern, either contraction, or expansion (shown). The remaining dots each have a randomly chosen direction component (unfilled). Arrows and filling shown here were not present in the dynamic stimulus and are added to assist the reader. The subjects nominated the location in which they perceived radial motion.

4.3 Experiment 1

4.3.1 Methods

The ability of participants to detect radially expanding and contracting patterns of optic flow was measured by establishing motion coherence thresholds in a spatial, two alternative forced choice (2AFC) task. In one, randomised, location an RDK comprised two populations of dots, totalling 75 dots in all (see figure 4.1). 'Signal' dots radiated from, or towards, the centre of an annulus of radius 5.8 degrees of visual angle. The centre blank region of the annulus was of radius 0.84 degrees hence the area of the window available for dots to be displayed in
was 52.8 degrees\(^2\). Dot density was therefore 1.42 dots per degree\(^2\). The speed of each dot was a linear function of its distance from the centre, introducing a speed gradient to the global pattern consistent with a flat object moving smoothly in depth (NB, although this relationship is not strictly correct, the approximation is adequate at small retinal sizes). The other population of 'noise' dots flowed in randomly chosen directions, presenting incoherent motions, and their speeds were similarly scaled to their eccentricity. The alternative location contained 75 noise dots only, i.e. all dots moved in random directions, their speeds being scaled with eccentricity matching the signal animation in terms of speed.

Using the method of constant stimuli the proportion of signal dots in the animation containing the radial pattern was varied to establish psychometric functions from which detection thresholds were estimated. (The total number of dots was maintained by adjusting the quantity of noise dots accordingly). Three observers took part, two experienced in psychophysics experiments, one of which was the author. In the case of the experienced observers (PB & RW) six levels of signal ratio were used to establish the psychometric functions, ranging linearly between 0.02\% and 0.42\%. Observer SH required greater signal strength to produce a complete psychometric function: 0.1\% to 0.58\%.

A range of dots (Difference of Gaussian) of differing centre frequencies (0.93 to 11.2 cycles/deg) were randomly distributed throughout the animation window, removing any texture gradient cues. Each animation comprised 25 unique movie frames, each of which was shown three times in succession before the next frame was displayed. The frame rate was 75Hz.

Three subjects participated in the experiment, all having normal uncorrected vision. The animation windows were centred 6.17 degrees to the left and right of fixation.
Independent Variables

Three variables were manipulated in the experiment. The ratio of coherently moving dots in the signal animation was varied to establish detection thresholds. In addition to this the spatial frequency change between the first and second presentations of each DoG element was varied (both noise and signal elements were manipulated in similar fashion). The nominal speed governing the range of individual dot speeds in the animation was also varied.

Spatial Frequency Manipulation

When an observer fixates the centre of a flat textured surface approaching at constant velocity, the change in size of each texture element is a linear function of its eccentricity from the centre of expansion (for small retinal angles). This will be termed the 'reference' simulation. The size and hence centre spatial frequency of each DoG element was manipulated with reference to the rate of expansion proscribed by its position relative to the centre of the annulus. The minimum spatial frequency was 1.65 cycles/degree, the maximum 14.7 cycles/degree. Centre frequencies were calculated using the function described in Chapter 1 (Clement, 1993).

Four conditions of rate of element expansion were explored: zero growth, veridical growth, twice the veridical rate, and four times the veridical rate. Figure 4.2 indicates how the size of a DoG changes between consecutive samples of its trajectory. Figure 4.3a shows the 'reference' simulation of motion-in-depth, with element size change a function of element divergence. Figure 4.3b shows a motion-in-depth simulation where elements grow at a greater rate than divergence rates specify.
Figure 4.2 Element size changes of Difference of Gaussian (DoG) elements. A veridical simulation of an approaching object increases in retinal size between first (left) and second (right) samples of its motion. The growth rate of elements was manipulated independently of the rate at which elements diverged from each other. Panel A
illustrates zero growth rate between samples. Subsequent panels show veridical growth (B), twice the veridical rate (C) and four times the veridical rate (D).

Figure 4.3a. Simulating motion in depth. The schematic shows the geometry of the ‘reference’ motion in depth simulation, i.e. dots distributed on a fronto-parallel plane, are to ‘approach’ the observer. The dots at point A in the upper panel, at time 1, are to simulate moving to point B, lower panel at time 2. To calculate their image size and spatial relationships in the original plane (the computer display), project from the observation point through the new position. The simulated motion-in-depth has preserved the relationships between dot size and the centre-to-centre distance between texture elements (divergence).
Figure 4.3b. The geometry of simulating a field of two random dots (black rectangles) approaching an observer, with dot-size increasing at a greater rate than specified by divergence. The dots at point A in the upper panel (at time 1) are to move to point B, lower panel at time 2. To calculate their spatial relationships and size in the original plane, project from the observation point. To distort element growth rate the size of the dots in the projection plane has been exaggerated. The centre-to-centre relationships between elements has been preserved, keeping divergence rates among dots equal to the ‘reference’ simulation depicted in figure 4.3a.

**Speed Manipulation**

Three conditions of nominal speed were explored. As the speed of each element was dependent on its location in the annulus the nominal speed was specified in terms of the speed at the periphery, all other dots being slower than this according to their distance from the centre of the window. Because of this speed gradient the speed for each condition is quoted at half the radius, namely: 1, 2 and 4 degrees/s.

Both speed and growth manipulations affect the spatial frequency change during the animation. Since an animation presents linear time slices of a hypothetical
continuous motion-in-depth a high speed results in larger retinal size increments than a low speed. The role of the growth manipulation was to distort this attribute. To appreciate the consequences of both these variables on the spatial frequency change in the animation refer to Table 4.1. Table 4.1 shows the average spatial frequency change between any two frames of animation, for each combination of the speed and growth variables. It is readily apparent that the highest speed and most exaggerated growth rates produce the largest changes in size and spatial frequency between successive samples of a DoG’s trajectories.

<table>
<thead>
<tr>
<th>Element growth multiple = 0</th>
<th>Speed (deg/s)</th>
<th>2</th>
<th>Speed (deg/s)</th>
<th>4</th>
<th>Speed (deg/s)</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Element growth multiple = 1</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Element growth multiple = 2</td>
<td>2.14%</td>
<td>5.00%</td>
<td>10.18%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Element growth multiple = 4</td>
<td>4.88%</td>
<td>9.79%</td>
<td>17.85%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Element growth multiple = 8</td>
<td>9.59%</td>
<td>17.26%</td>
<td>28.35%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
</tbody>
</table>

Table 4.1 Mean change in the peak spatial frequency of RDK dots as a function of speed and element growth conditions.

Therefore the experiment comprised twelve unique conditions. Each condition was explored in a separate block of trials, and each participant performed at least four blocks of trials for each condition. Sixty observations were recorded in each block, resulting in at least 240 observations for each psychometric function generated.
Figure 4.4 Coherence thresholds for detecting radial motion. Three curves are shown, one for each speed condition explored. Thresholds on the y-axis are plotted as a function of the element growth rate, on the x-axis. Each data point represents the mean threshold of three observers. Error bars are one standard error of the mean.

4.3.2 Results

Figure 4.4 shows motion coherence thresholds as a function of element growth rate in multiples of the veridical rate. Three curves are shown, one for each speed condition. Blue diamonds represent a mean speed of 1 degrees/s, pink squares represent 2 degrees/s and red triangles represent 4 degrees/s. Each data point is the mean threshold of three observers, and the error bars are one standard error of the mean.

The most striking feature of the curves is the reduction of threshold in the zero growth condition, when elements were constrained not to change in size over their lifetime. A two way, repeated measures analysis of variance (ANOVA) was used to establish significant main effects of element growth ($F_{(3,6)}=4.99$, $p<0.05$), and speed ($F_{(2,4)}=13.39$, $p<0.05$), as well as a significant interaction between these factors ($F_{(6,12)}=3.16$, $p<0.05$). The effect of speed, and the interaction between speed and growth are due to the reduced thresholds obtained
when growth=(4*veridical). When speed=1 degree/s thresholds decrease, which looks slightly anomalous compared to the thresholds for the other two speed conditions when growth=4. Of more interest is the effect of growth, which is most pronounced between growth=0 and growth=1 (veridical). A distinct drop in thresholds is evident, showing that no change in element structure leads to better performance than veridical growth. At first the finding that zero growth rates are optimal for optic flow detection appears counter-intuitive. It might have been expected that the visual system would be most sensitive to element growth rates that matched divergence rates of the real world. If sensitivity were determined at the global level this might be so. However, prior processing of signals at local level might favour stimuli optimised for the spatio-temporal characteristics of scale-restricted simple motion detectors. These results support this interpretation, previously applied by Ledgeway (1996) and Kim & Turano (1999) in the case of the discrimination of the direction of wide field translation. Although it has been argued that the architecture of the visual system should reflect the features of naturally occurring motion signals (Attneave, 1954, Barlow, 1961), these results seem to suggest that these are not the only constraints that impose themselves on the design of the system. Eckert & Zeil (2001) discuss from an ecological viewpoint that there are many influences that determine an optimal system, each having its own implications for survival. The need to process signals efficiently may contribute in a way that appears to subvert the importance of worldly features such as localised size change, as might the need to achieve sparse representations, redundancy reduction or reliability.

There is an indication that at the highest growth condition, thresholds are rising, consistent with a breakdown in either local or global processing. Unfortunately it is not possible to say at this stage which level of processing is responsible for the breakdown, though the results of the next experiment will give an indication as to the source.
4.4 Experiment 2

Experiment 1 used radially symmetric difference of Gaussian (DoG) dots because they offer a stimulus of restricted spatial frequency content. Manipulating their size manipulated the spectral content of the stimuli. In that experiment it was found that sensitivity to radial patterns of global motion was highest when the dots were constrained to remain of fixed size (and hence spatial frequency), contrary to the naturally occurring increase in retinal size of a genuine object approaching in depth. Previous work (Regan & Hamstra, 1993; Regan & Vincent, 1995; Vincent & Regan, 1997; Beverley & Regan, 1983), has shown that veridical element growth is not necessary to support accurate estimates of time to contact, and when available play a minor role except where they are in gross conflict with other motion-in-depth cues (Harris & Giachritsis, 2000). Early studies have characterised the effect of element growth as a matter of size, rather than of changing spatial frequency content.

The results of Experiment 1 do not reveal whether the increased sensitivity found in the zero growth rate condition was due to shared element size, or shared spatial frequency content. At first sight the issue may seem unimportant, as these two characteristics normally co-vary. However, for this experiment a stimulus has been designed which allows the manipulation of dot size while maintaining spatial frequency content constant (see Methods section). With this new stimulus Experiment 2 measured sensitivity to detect radial optic flow across a range of speeds and element sizes, as before, using the motion coherence method to establish the sensitivity of subjects.
Luminance plots in two spatial dimensions (x and y) showing the effect of filtering a binary dot. Refer to Panel A. Convolving a dot (left) with a high frequency filter (centre) results in an isometric dot of defined spatial frequency content. This will be referred to as a ‘circular grating’. Panels B and C show how the same filter applied to different dots results in spectrally similar dots, differing in size.

4.4.1 Methods
As for Experiment 1, with the design of each dot being the only difference. To produce differing dot sizes with constant spatial content a two-dimensional bandpass filter (of Gaussian profile in the spatial frequency domain) was applied to a conventional circular dot, i.e. one whose luminance profile contains only
two grey levels representing black and white (see figure 4.5, panel A, left). The effect of applying a filter (figure 4.5, panel A, centre) was to create a circularly symmetric circular grating of defined spatial content (figure 4.5, panel A, right). (NB consequently the spectral content of these dots was different to those used in the first experiment). By choosing different sizes of pre-filtered dot the diameter of the resulting circular grating was manipulated, independently of spatial frequency. The Gaussian-shaped filter (in frequency space) is defined by the following formula:

\[ y = \frac{1}{\sqrt{2\pi} \sigma} \ast e^{-\left(\frac{(r-f_{pk})^2}{2\sigma^2}\right)} \]

where,

- \( \sigma \) = one standard deviation of the filter 0.084 degrees (at 57cm)
- \( r \) = distance from the centre of the circular filter
- \( f_{pk} \) = the peak frequency of the filter, 0.29 degrees (at 57cm)

Applying this filter to each of a range of dots resulted in isometric bandpass circular gratings of average centre frequency 1.44 cycles per degree (at the viewing distance of 57cm) and average bandwidth of 1.8 octaves. Figure 4.6 shows the power spectrum of a sample of dot sizes after filtering. It can be seen that the spectra are not identical for all filtered dots, however the 'formants' (peaks and troughs) are a consequence of the discrete Fourier transform which can only return frequencies of integer multiples. These do not reflect the responses of frequency selective cells found in mammalian vision that are known to integrate over comparable ranges of frequencies (e.g. 1-2.5 octaves,
Georgeson, 1980; Wilson, McFarlane, & Philips, 1983), and which would consequently respond similarly to each of the amplitude profiles shown in figure 4.6. Centre frequency and bandwidth were estimated for each dot size and the figures quoted above represent the average for the set. The range of sizes of dots (pre-filtering) was: 0.084-0.29 degrees.

![DoG frequency spectrum](image)

Figure 4.6 Frequency spectra of a sample of the filtered dots of Experiment 2. Dots of different sizes were convolved with the same filter, resulting in isotropic elements differing in size, yet sharing a common spatial frequency composition. Sizes of dots prior to filtering are listed in the insert.

### 4.4.2 Results

Figure 4.7 shows motion coherence threshold as a function of element growth rate in multiples of the veridical rate. Three curves are shown, one for each speed condition. As before, blue diamonds represent a mean speed of 1 degree/s, pink squares represent 2 degrees/s, and red triangles represent 4 degrees/s. Each data point is the mean threshold of three observers, and the error bars are one...
standard error of the mean. The graph clearly shows an effect of speed, with motion coherence thresholds elevated in the 4 degrees/s condition. This was confirmed to be a significant main effect of speed with a two way, repeated measures ANOVA ($F_{(2,4)}=12.36, p<0.05$). The analysis of variance revealed no reliable effect of element growth ($F_{(3,6)}=1.78, p>0.05$), and no significant interaction between the two factors ($F_{(6,12)}=0.354, p>0.05$). A Bonferroni/Dunn post hoc test showed that the difference was between the slow (1 degree/s) and fast (4 degrees/s) conditions ($p<0.01$).

There are two points to note from these results. Firstly, the sensitivity advantage previously obtained in the zero growth rate condition has been abolished. However, this is not to say that sensitivity in this condition has declined, rather that thresholds remain low under other conditions. This is most pronounced in the slow speed (1 degree/sec) curve, where thresholds of around 17% are comparable with optimum thresholds found in Experiment 1, zero growth condition. The second point to stress is the pronounced effect of speed on thresholds. A clear drop in sensitivity accompanies increasing speeds. The relative absence of this effect in the previous experiment indicates that controlling spatial content has produced this phenomenon. This point will be elaborated in the Discussion.
Figure 4.7 Coherence thresholds for detecting radial motion. Three curves are shown, one for each speed condition explored. Thresholds on the y-axis are plotted as a function of the element growth rate, on the x-axis. Each data point represents the mean threshold of three observers. Error bars are one standard error of the mean.

4.5 Discussion

Experiment 1 examined the effect of element growth on the ability to detect optic flow. The rate of change of the size of dots in a radially expanding RDK pattern was manipulated, independently of the rate at which dots diverged. Sensitivity was found to be greatest when consecutive samples of dots’ trajectories were identical in size. However, as size and spectral content co-varied it was not possible to tell whether size or spatial frequency content was responsible for the enhanced performance. If spatial frequency was the cause of the effect, this could be attributed to the characteristics of low-level local motion detectors. If size were the crucial factor this would locate the cause of the effect at a higher level, perhaps a mechanism addressing the correspondence problem, globally. To tease size and spectral content apart, Experiment 2 presented a similar radially expanding field of dots with the crucial difference that dot size could be manipulated without drastically altering its spatial frequency signature.
Sensitivity to detect expansion was found not to vary with the rate of size change of dot elements, implying that the credit for enhanced detection in Experiment 1 belongs to the presence of common spatial frequencies, rather than common size. Much previous work on the perception of motion-in-depth has characterised the stimuli used in terms of size, and size change, rather than in terms of spectral content. This has been true in recent studies using random dot stimuli (Harris & Giachritsis, 2000), or earlier work using geometric figures such as textured squares (e.g. Regan & Hamstra, 1993; Regan & Vincent, 1995; Vincent & Regan, 1997). The use of the term 'size' may have obscured the role played by early motion systems in the performance of the tasks set. Size is, after all an 'object level' property, and dependence on its change for the perception of motion-in-depth would require a parsing of the scene into its discrete entities prior to analysis of the type of motion present. Referring to the 'size' of stimuli conceals at least two other means by which an approaching object may be perceived as such. A looming square with a uniform surface contains relative motion between its edges, analogous to the divergence among random dots seen in a radially expanding RDK. Indeed Freeman & Harris (1992) suggested a name for the psychological mechanism responsible for perceiving looming: the Relative Motion System. Another candidate for registering looming, recently reported by Schraeter, Knill & Simoncelli (2000) is that scale change alone can signal motion-in-depth. These authors found that a looming percept could be induced without any net motion signals (in the conventional sense) being present in their display. Their animation comprised several frames of independently generated white noise, each of which was passed through a lower frequency filter than the last. This resulted in a sequence of frames that were uncorrelated, but when shown sequentially presented a smooth transition along the spatial frequency spectrum from high frequencies to low, i.e. a progressive reduction in spatial scale. Observers were able to match this sequence to an optic flow animation comprising radially expanding dots (of fixed size) implying that the perception of looming may in part be mediated by a progressive reduction in the spatial scale of the image. Observers also experienced motion after-effects to the
changing scale sequences, suggesting that they do indeed stimulate motion mechanisms. Schrater, Knill & Simoncelli (2000) ensured that conventional local motion detection mechanisms were not responsible for the perception of expansion by using uncorrelated frames of noise. Still, their result is controversial following a recent failure to replicate (Rogers & Anstis, 2002). In view of this the results obtained in the experiments reported here have been interpreted to reflect processing limitations at the local level. This is in line with the conventional view that the properties of local motion detectors place a constraint on the capabilities of global motion detectors (e.g. Ledgeway, 1996), and supports the hierarchical motion processing architecture proposed by Maunsell & van Essen (1983) and Van Essen & Maunsell (1983).

A second feature of the results of Experiment 2 was a clear effect of speed, which was not apparent in Experiment 1. Higher speed in apparent motion sequences results in a larger jump between first and second samples of a dot’s trajectory. In Experiment 1 this increased jump was accompanied by a change in spectral content, and no performance penalty. In Experiment 2 spatial frequency remained the same, and performance decreased. This implies that local motion detectors prefer frequency change to occur across large distances. In the past apparent motion has been characterised in terms of two systems each operating over a different spatial extent (Braddick, 1974), the so-called ‘short-range’ and ‘long-range’ processes. Anderson & Burr (1987) later suggested that motion detection operates over a continuum of spatial scales due to the spatial frequency specificity of motion detectors. In order to respond to features at a specific scale the receptive field size of a detector must necessarily be large. Combining, minimally, two feature-detectors to sample motion across space, implies that low frequency motion detection would operate over a greater distance than analysis at high spatial scale. As pointed out earlier, when observing approaching texture a greater lateral distance across the retina is accompanied by a size change, making the second of these feature-detector components more likely to encounter a larger, or lower frequency, component. A tentative conclusion from the data is that at high speeds the lack of spatial frequency change meant that
second-place feature-detectors were deprived of their optimal input, therefore leading to a drop in detection performance.
Chapter 5
Nulling the Radial Motion Speed Illusion

5.1 Abstract
Independent laboratories have confirmed that the global speed of a radially expanding pattern is perceived to be greater than that of a rotational pattern that is matched for local speeds. Three experiments are reported that investigated the source of this speed illusion. Conflicting depth cues in the form of static disparity, dynamic disparity and element size change were introduced in an attempt to diminish the magnitude of the illusion. None of these manipulations was effective in compromising the exaggerated perception of speed. These results have implications for models of cue combination, which should reflect the lack of impact of disparity cues on perception in the conditions of relative motion explored here.
5.2 Introduction

Several studies have shown that a radial motion pattern is perceived to move faster than a rotational motion pattern even when they are matched for local element speeds and directions (see figure 5.1). The magnitude of the effect varies greatly between studies and values between 10% and 60% have been reported. The phenomenon has been found to be present across a range of dot speeds (Bex & Makous, 1997; Clifford, Beardsley & Vaina, 1999), and is obtained whether the stimuli comprise spatially narrowband gratings (Bex & Makous, 1997; Bex Metha & Makous, 1998; Clifford, Arnold & Wenderoth, 2000) or broadband dots (Clifford, Beardsley & Vaina, 1999; Geesaman & Qian, 1996; Geesaman & Qian, 1998). Also, it is found regardless of the presence of speed gradients in the patterns (Clifford, Beardsley & Vaina, 1999; Watson & Bex, 1999).

(NB the term ‘speed illusion’ has been applied to the perceived speed of radial patterns because prior research has shown that rotation and translation patterns are judged to have the same speed when the retinal speeds of individual elements making up their flow fields are matched (Geesaman & Qian, 1996, 1998). Only with radial patterns do the same retinal speeds produce a different perceived speed, so by convention the existing literature has treated this pattern as the unusual case. The use of the term ‘illusion’ is not intended to imply that some patterns are perceived ‘veridically’ and some not).

When rotational and radial stimuli are carefully matched overall for speed and directional content, it is evident that the difference in the perceived speeds is not due to the local structure of the patterns, which has an identical distribution of speeds in both cases, i.e. maximum, minimum, mean, median and mode are the same. Retinally then, these patterns are equated, and a simple comparison on the basis of retinal speed signals would reveal no difference between them. The fact that observers in such experiments consistently choose the radial pattern as appearing faster, when in fact it is equal, shows that a more abstract measure of speed is being perceived, even though it is not clear at this stage what that measure represents. Since the only difference between patterns is the spatial
arrangement of the motion vectors this must be the attribute that is responsible for the dissimilar estimates of speed. This appreciation of the global structure is consistent with that needed to interpret the optic flow generated when an observer and their environment move relative to each other (Gibson, 1950). It has been suggested that the speed perceived with radial patterns is that of the implied motion in depth, rather than the speed of the dots themselves (Clifford, Beardsley & Vaina, 1999). i.e. "We perceive the motions of objects, not the motions of retinal image features." (Croner & Albright, 1999). Clifford et. al., (1999) termed this the 'motion-in-depth' hypothesis. It was decided to test this hypothesis by exploiting additional sources of depth information to alter the perceptual interpretation of the radial motion pattern, and hence influence the magnitude of the speed illusion.

Figure 5.1 Example of the stimuli used in a typical speed matching experiment. An animation is shown in each of two temporal or spatial locations. In one location a field of dots rotates about the centre of a window (left panel). Their direction could be clockwise or counter-clockwise. In the alternative location dots radiate relative to the window's centre, producing an expanding or contracting pattern. To ensure patterns are equated in speed and direction a radial pattern is created, then a 90 degree directional offset is added to each dots directional parameter to produce rotation. Since the same process is used to generate both patterns, and the 90 degree offset is a constant, both
patterns are equated in terms of speed and direction. The observer's task is to nominate which location appeared the faster.

**Relative Motion, the Kinetic Depth Effect (KDE) and Motion Parallax**

Optic flow is frequently simulated on the computer monitor using the Random Dot Kinematogram (RDK). This comprises an animated field of dots on an otherwise empty background that gives the impression of structured motion without the complicating addition of objects or surfaces (Nakayama & Tyler, 1980). According to the motion-in-depth hypothesis a radially expanding RDK provides a strong impression of depth due to the relative motions of the dots making up the pattern. The revelation of shape through object motion (e.g. rotation) was coined the kinetic depth effect (KDE) by Wallach & O'Connell, (1953), often also termed 'structure from motion'. Relative motion between an observer and their surroundings can similarly reveal an impression of depth structure, and is named motion parallax. Both concepts are closely related, so the more general term relative motion (RM), will be used.

Although the perception of depth is difficult to measure quantitatively, Ringach et al., (1996) monitored eye movements to a rotating sphere simulated using random dots, finding that vergence angles mirrored those expected of subjects tracking the surface of a real spherical object. Many subjective studies present converging evidence that KDE gives a compelling impression of three-dimensional structure (e.g. Dosher et al, 1989; Todd et al, 1988).

In an expanding radial motion pattern the dots in an RDK radiate away from the pattern's centre, becoming progressively more distant from one another as they proceed to greater eccentricities. The relative motion of the dots gives observers a strong sense of either a) a flat textured surface approaching in depth, or b) travelling through a volume of individual elements, each at a different depth. In life, each scenario is associated with its own characteristic speed gradient. (If the observer looks in the centre of the pattern, the focus of expansion, the speed of each element is a function of its radial distance from the point of fixation. A gradient of speeds is therefore found along any radial arm emanating from
fixation). The same is true for the rate at which individual dots grow in size as they 'approach' the viewer-each scenario has characteristic parameters. However, it has been found that speed gradients (Freeman, Harris & Tyler, 1994) and element growth rates (Regan & Hamstra, 1993; Regan & Vincent, 1995) are not necessary to support accurate estimates of time to contact, and when available play a minor role (Harris & Giachritsis, 2000) except where they are in gross conflict with motion cues (Harris & Giachritsis, 2000, see their Experiment 3). In the case of element growth this is fortunate when generating a rotational pattern that matches a radial pattern in all local velocity vectors. By applying either a 90 or -90 degree offset to the direction of each dot's travel the radial pattern is transformed into a rotation. If the dots in each animation are set to maintain a constant size adequate faximilies of both types of optic flow are generated.

*Binocular Disparity as a Depth Cue*

Depth information is lost when the visual array is projected onto the two dimensional retina. It is recovered by the brain as a consequence of the slightly differing views that the eyes have of the world, which is a function of their lateral displacement. When a point in space is fixated the eyes rotate to centre the area of attention on the fovea. The fixated region falls on corresponding points on the retina, i.e. geographically equivalent tissue on each retina is stimulated. Objects placed at differing depths do not fall on corresponding points and the degree of mismatch between the sites of stimulation is exploited by the brain to assess the depth of the non-fixated objects relative to the plane of fixation. The mismatch is referred to as disparity, and the geometry that gives rise to this relative depth cue can be appreciated by referring to figure 5.2.
Figure 5.2 The geometry of crossed (panel A) and uncrossed (panel B) disparities. Point F is the point of fixation, point O is the location of an object in the scene. In panel A the object is closer than the fixation plane, and its projection falls on opposite sides of the fixation points on the retinae, as indicated by the directions of the arrows. Panel B shows how an object further from fixation projects to the same side, relative to the fixation image.

When an object moves along the z-axis (motion-in-depth) relative to a fixated point, its disparity changes over time, producing dynamic disparity. Ordinarily objects moving in depth present monocular as well as binocular cues that reflect the direction of progress. However, it has been shown many times (Julesz, 1971; Norcia & Tyler, 1984; Regan, 1993; Cumming & Parker, 1994) that disparity alone contains sufficient information to yield a percept of motion in the third dimension using dynamic random dot stereograms (DRDS). In a conventional motion sequence consecutive frames contain similar images that are seen to move in depth. However motion in depth can be generated without such luminance correlation (a ‘first order’ cue) between frames if a new image is generated for each stereo-pair of the animation. The required degree of disparity can still be introduced to each stereo pair producing a sensation of motion in depth at a speed that is proportional to the rate of change of disparity (Cumming & Parker, 1994) yet containing no object motion. Several studies attest to the veracity of this second-order motion cue, including work on detection thresholds (Cumming & Parker, 1994; Gray & Regan, 1996, Westheimer, 1990), and speed discrimination thresholds (Brooks & Mather, 2000; Portfors-Yeoman & Regan, 1996; Portfors & Regan, 1997). Refer to Patterson (1999) for a review. In the experiments reported here both static disparity (Experiment 1) and dynamic disparity (Experiments 2 & 3) will be introduced to imply fixed depth or motion-in-depth, respectively.
Cue Disagreement: Interactions between Relative Motion and Binocular Disparity

In the following three studies by Regan and colleagues the changing size of their stimuli (geometric figures) is considered to represent a relative motion cue, albeit a less dense one than is normally provided by a cloud of dots in an RDK stimulus.

Gray & Regan (1996), presented subjects with a dot oscillating in size, which is reported to have yielded a compelling sensation of motion-in-depth. Using the method of adjustment the authors found that this sensation was cancelled when the disparity of the dot was oscillated in the opposite sense, i.e. to imply an approaching dot when its size was diminishing, and vice versa. Unfortunately no quantitative metric was used to record the effect. A further precedent for expecting an interaction between RM and dynamic disparity comes from Regan & Beverley (1979). By exposing subjects to an expanding square for a period of 20 minutes they induced a motion after-effect (MAE) in a static square test pattern that was perceived to recede in depth. (In the adaptation phase the square would expand from small to large, disappear, then the cycle would repeat). They found that this after-effect could be nulled by the application of an appropriate quantity of binocular disparity.

Gray & Regan (1998) collected error data on the ability of observers to estimate the time to collision (TTC) of a dot stimulus, which could increase in size (regarded as relative motion between points around the circumference for present purposes) and/or change in binocular disparity. Errors when both cues were congruent were lower than when one or the other was inappropriate, e.g. if disparity signalled approach, the dot did not grow in size; if dot size increased then disparity was held constant.

In short, then, the literature on depth from relative motion and dynamic disparity suggests that both cues are robust in producing salient three-dimensional percepts, and that they interact when available simultaneously. If relative motion is indeed responsible for the 'misperception' of speed in a radially expanding
pattern compared to a matched rotating pattern it ought to be possible to present disparity cues to modify the magnitude of the speed illusion by implying alternative or conflicting events.

5.2.1 An Overview of the Experiments

In these experiments binocular disparity was introduced to radially expanding patterns to give additional information about their location in depth. In the first experiment static disparity was added to define the pattern as existing at one of several fixed depth planes.

Experiment 2 took a similar approach in presenting binocular depth cues. In this experiment, however, dynamic disparity cues were presented. Where the relative motion depth cue implied by a radially expanding stimulus suggests an approaching object, the disparity cues were arranged to conflict with this by simulating an object receding in depth. This was achieved by varying the disparity from crossed at the beginning of the animation to uncrossed at the end (see figure 5.2). If the motion-in-depth hypothesis is correct it is anticipated that such conflicting information would null the speed illusion, similar to the MAE nulling found by Regan & Beverley (1979).

Experiment 3. In Experiment 2 the size of individual dots in the RDK remained fixed, a stimulus attribute that is not consistent with either approaching or receding interpretations. Experiment 3 manipulated the dot size to be consistent with a receding object, thus presenting an additional conflicting depth cue in a further attempt to null the speed illusion.

5.3 Experiment 1. Fixed Disparity

In previous experiments the speed illusion has been measured by observing a single expansion field (as opposed to a stereo pair). Therefore disparities were always zero (where viewing was binocular), consistent with a fixed depth plane. (Radial expansion in a fixed depth plane is consistent with a deformation event, for example an inflating balloon). To provide benchmark figures for the magnitude of the illusion with zero and non-zero disparities its strength at a
range of depth planes was measured, similar to those that would be present in later experiments.

5.3.1 Methods

Observers
Three subjects took part in the study, all of whom were chosen to be naive to the aims of the experiment because of the subjective nature of the judgement they had to make. Each had normal, or corrected to normal visual acuity and good stereo-acuity.

Each observer performed at least 4 (observer MC), more usually 6 (observers AC & SK), blocks of trials, comprising 40 trials per block.

The ability of the volunteers to discriminate motion-in-depth was confirmed in a separate experiment. A forced choice procedure was used to establish if subjects could tell when a dynamic disparity RDK 'receded' (crossed disparity of 20 pixels to uncrossed disparity of 20 pixels) or 'approached' (The dots were static, i.e. there was no relative motion among them). Two subjects scored better than 95% correct on this task, while SK performed much better than chance at 76% correct. While this lower performance suggests poorer stereo-motion perception for SK the subsequent pattern of experimental results for this observer was similar to that of the others.

Stimuli
The dots (Difference of Gaussian) used had a centre spatial frequency of 1.6 cycles per degree of visual angle at the viewing distance of 1 meter employed. Each dot subtended approximately 0.573 degrees of visual angle. One hundred dots occupied each window, which was curtained at a radius of 3.33 degrees by a raised cosine function. The dot density was therefore 4.8 dots per degree².

Each dot had a long lifetime of 20 frames to aid the stereo percept.
The ‘standard’ radial pattern moved at a fixed speed at all times, nominally 2.98 degrees/s. (As the speed of each dot was scaled according to its eccentricity from the centre of the dot field to produce smooth expansion the median speed of the dots was 0.997 degrees/s. Therefore dots moving at 2.98 degrees/s were only seen at the outer extremity of the window). The speed of dots in the rotation pattern was varied according to a Quest algorithm (Watson & Pelli, 1983). Quest selected speeds from the range: \(0.298 \ldots 26.26\) degrees/s in \(\log_{10}\) steps. The speed of dots in the rotation pattern was similarly scaled with eccentricity to generate a rigid rotation that was locally matched to the radial pattern.

Stimuli were viewed through a Wheatstone mirror stereoscope to allow the introduction of binocular disparity to the image. To this end each movie actually comprised two windows of animation with similar, but non-identical images (see figure 5.3). Each field of dots was displayed within a rectangular black border that served as a reference frame, found to be effective in producing a strong perception of motion-in-depth (Erkelens & Collewijn, 1985). This frame was presented prior to the commencement of the experimental trials, and subjects were required to fuse the rectangular frames while viewing the display through the stereoscope. (Pre-trial fusion has been shown to improve the formation of a stereoscopic percept (Uttal, Davis & Welke (1994)). The front surface mirrors were arranged so that only the appropriate monocular image was seen by each eye. Observers were asked to fixate a small black line located midway along the top border of the reference frame. The animation was therefore centred 2.86 degrees from fixation. Once fusion was obtained the subject started the block of trials by pressing one of the response keys used to record their judgements: buttons 1 and 3 of the keyboard’s numeric keypad. The two fields of dots within each rectangle were identical to each other, but were placed in different locations relative to their reference frames. On a given trial each field of dots was displaced up to 30 pixels from the centre of their reference frames. The direction of this displacement in one half image was opposite to that of its partner, giving a maximum of 1.43 degrees of relative disparity. The displacements used were: 0, 10, 20, 30 pixels. Each of these displacements (except 0) was used to produce
a crossed and an uncrossed disparity resulting in seven unique conditions of relative disparity: 0, ±0.238, ±0.477 and ±0.714 degrees of visual angle. The metric distances in depth implied by these changes were as follows. For crossed disparities, the dots appeared to float at a depth of (12.2, 21.7, 29.4) cm closer to the observer than the fixation plane. For uncrossed disparities the values were (16.1, 38.5, 71.5) cm behind the fixation plane.

Procedure
Participants were asked to discriminate between two complex motion patterns on the basis of their apparent speed, their task being to nominate which animation appeared faster. (No instruction was given regarding which dimension of motion, lateral or in depth, should be judged—observers were allowed to select their own criteria). Two temporal intervals each contained a 1000mS movie depicting a field of dots travelling in either a radial or a rotational motion. The radial pattern was always expanding, however the rotating pattern was varied randomly between clockwise and anti-clockwise directions. The inter-stimulus interval was approximately 1000mS, during which the next animation was calculated. Radial motion was the ‘fixed’ animation, i.e. its speed did not vary.

Between trials the reference frame and fixation mark were continuously displayed to allow participants to maintain fusion. Subsequent trials were triggered once the observer’s response was made.
Figure 5.3 A stereo pair showing an expanding stimulus. Each image shows a field of dots travelling in the directions indicated by the arrows. By viewing through a Wheatstone stereoscope, only one image is seen by each eye. Each image contains exactly the same spatial arrangement of dots, but those in the right panel are displaced to the right relative to the surrounding black frame. The mark on the top edge of the windows is the point that subjects were asked to fixate during a trial.

5.3.2 Results

Psychometric functions were fitted with a cumulative Gaussian function where the mean and standard deviation of the function were varied to obtain a least squares optimisation between the data and the fitted curve. The point of subjective equality (PSE) was estimated from the fitted curve by calculating the inverse of the cumulative Gaussian at the point where subjects judged the radial pattern faster with a probability of 50%.
Figure 5.4 Results of Experiment 1. The magnitude of the speed illusion is shown for three observers separately (figures 5.4A, B, C). Plotted on the y-axis is the speed of the dots of a rotating pattern (as a factor of the radial speed) that provided a perceptual match to the (fixed speed) radial pattern. The speed illusion is shown as a function of the amount of fixed horizontal disparity on the x-axis, and is shown separately for each of the three observers. The two curves shown on each graph are for crossed and uncrossed disparities, and results for zero disparity are plotted separately on the figure, see legend. The horizontal line shows the dot speed expected if subjects had perceived a match (in retinal speed) terms between rotation and radial patterns.

Figure 5.4 summarises the data from the three observers. Each graph shows data for one subject, plotting the point of subjective equality (PSE) on the y-axis as a function of the relative disparity condition on the x-axis. The two curves on each graph represent crossed disparity (nearer than the plane of fixation) and uncrossed disparity. The solid line shown for reference is the retinal speed match, i.e. the PSE that would be obtained if subjects saw a rotational pattern with nominal mean dot speed of 0.997 degrees/s as the same global speed as the radial pattern (whose mean dot speed was always 0.997 degrees/s). Values above this line indicate that subjects needed a faster rotation to provide a perceptual match to the radial pattern.

It can be seen from the graphs that in the zero disparity condition the speed illusion is evident, and consistent with previous studies where disparities were zero due to viewing of a single pattern of motion (not a stereo pair). Comparing between subjects, first, it is apparent that the speed illusion is subject to individual differences. Subject MC requires in excess of a 60 % increase in the dot speed of a rotation pattern to match the expanding pattern. Subjects AC and SK require a greater increase of over 100%. These biases are consistent with previous work, but are larger than reported elsewhere, possibly reflecting the effect of stereoscopic viewing, which has not previously been explored.
The zero disparity condition acts as a benchmark against which to compare performance to non-zero disparities. The data show no evidence that fixing the stimulus at constant depths using non-zero disparity affects the apparent speed exaggeration. Despite the addition of up to 1.43 degrees of relative disparity, there is no sign of a consistent change in the speed illusion in any of the subjects. This was confirmed statistically with a repeated measures, one way analysis of variance, $F_{(5,12)}=0.981$, $p>0.05$.

In summary, placing the RDK's dot field at a constant depth does not diminish the size of the speed illusion. Therefore, any effect found in later experiments is not due to an artefact of the viewing conditions necessary to induce stereoscopic fusion.

5.4 Experiment 2. Dynamic Disparity

As outlined in the introduction, few studies have pitted relative motion against dynamic binocular disparity in a stimulus simulating motion-in-depth. Of those that have, Gray & Regan (1998) is the most relevant here, assessing the effect of disparity on the estimated time to contact of an expanding dot. Implicit in their observers' judgements is the perception of motion-in-depth, and the ability of disparity to modify its magnitude implies that cue interaction can be expected. However, their results do not reveal whether the speed illusion is a consequence of the sensation of motion-in-depth. Therefore, in the following experiment dynamic disparity was introduced, signalling a receding dot field, while the relative motion in the radially expanding pattern is presumed to signal an advancing textured surface.

5.4.1 Methods

Methods were the same as for Experiment 1, with the exception of the disparity manipulation. Whereas the previous experiment presented a fixed disparity throughout a trial, disparity in Experiment 2 varied during the trial. On the first frame of the animation the dots had 0.955 degrees (2*20 pixels, 20 per eye) of crossed disparity. On each subsequent frame the disparity was altered in linear
steps to end up with 0.955 degrees (2*20 pixels, 20 per eye) of uncrossed disparity on the final frame. This had the effect of making the pattern appear to move smoothly in depth, starting nearer to the observer than the plane of the fixation rectangle, and ending further away than it. The extremes of the disparity range were chosen to be 0.955 degrees as fusion was difficult to obtain at greater values. The total metric distance in depth implied by these changes was 60.2cm. The crossed condition was consistent with a target 21.7cm closer to the observer than the fixation plane, and the uncrossed condition signalled a target 38.5cm behind it. The resulting speed of motion-in-depth was 60.2 cm/s as the stimulus duration was 1000mS. The retinal speed of each field of dots (due to changing disparity) was 0.955 degrees/s. The relative speed of the dots in rotation and radiation were as described for Experiment 1.

5.4.2 Results

Points of subjective equality were calculated in the same way as for Experiment 1: by fitting a cumulative Gaussian to each subject's psychometric function, and calculating the hypothetical dot speed (of the rotation pattern) that corresponded to the subject claiming that radial motion appeared faster 50% of the time. Figure 5.5 shows data for each subject separately. The speed illusion has been plotted as a ratio for ease of comparison between differing speed conditions. Error bars indicate 95% confidence intervals. Filled bars indicate fixed disparity and are re-plotted from Experiment 1, zero disparity condition. This is the benchmark against which to compare the effect of conflicting dynamic disparity, shown by open bars. Clearly there has been no effect of dynamic disparity, as the illusion’s magnitude is undiminished in comparison with Experiment 1. This result was confirmed statistically using a repeated-measures ANOVA that found no significant difference between fixed disparity and dynamic disparity conditions (F(1,2) =0.411, p>0.05 ).
5.5 Experiment 3. Dynamic Disparity and Changing Size

Past studies have documented the effects incongruent size change on observers' responses to an approaching object, both real and simulated. For example, Savelsbergh et. al., (1991) reported that when an approaching ball was deflated in-flight, grasping movements were delayed, as might be expected if the size change influenced estimated time to contact (TTC). Harris & Giachritis (2000) found that shrinking element size in a simulation of an approaching pattern had a strong influence on TTC estimates, both accuracy and consistency being severely compromised. With these results in mind, and with the prospect that two conflicting cues might prove more effective when combined, element size was manipulated. Dots shrank during the course of the animation, implying a
receding object, in concert with the dynamic disparity cue. (Size change has the added advantage of being a 'first-order' cue, unlike disparity).

5.5.1 Methods
Methods were the same as for Experiment 2, with the added manipulation of dot size. As before, dynamic disparity was used to make the field of dots appear to recede in depth by starting the animation with 20 pixels of crossed disparity, proceeding to 20 pixels of uncrossed disparity by the end of the interval. As the dots receded in depth they were made to shrink in size, as a real object or texture element might do. Dots started off at the nominal size of 0.573 degrees, and shrank at a linear rate to 0.307 degrees by the end of the animation. Although real objects would decrease in size at an exponential rate the linear approximation over a short time period is not considered sufficiently different to affect the general impression of a surface receding smoothly in depth.

5.5.2 Results
Figure 5.6 shows the magnitude of the speed illusion for all three observers, expressed as a factor of the standard speed. For each observer data from Experiments 1 and 2 have been re-plotted for comparison. Filled black bars show the strength of the illusion when disparity was fixed at zero degrees (from Experiment 1). Open bars show the effect of dynamic disparity signalling a receding pattern (from Experiment 2). Finally, filled grey bars show the effect of a 'receding' dynamic disparity in conjunction with diminishing dot size. Error bars indicate 95% confidence intervals. A repeated measures ANOVA was conducted, comparing the fixed disparity condition with the results from Experiments 2 & 3. This showed that there was no significant effect of the combined disparity and size change cues on the magnitude of the speed illusion ($F_{(2,4)} = 1.68, p>0.05$).

Given previous demonstrations of the influence on TTC of changing element size, it is somewhat surprising that no effect on the speed illusion was found.
However, the effects found in past studies have been quite small, compared to the influence of the relative motion cues present. In addition, TTC is chosen as a dependent variable because it has the advantage of being independent of any estimate of absolute depth. Regan & Hamstra (1993) have shown that TTC can be judged solely by the ratio of stimulus size to rate of change of stimulus size, knowledge of depth being unnecessary. Speed estimates in terms of distance covered in unit time require an absolute depth estimate, something that cannot be reliably discerned from the stimuli employed here. It may simply be that using apparent speed is not an adequate measure to register the small effects of discrepant size change noted elsewhere. Even so, it is noteworthy that when size and disparity contradict relative motion they are not capable of extinguishing or even diminishing the strength of the illusion.

![Results from Expt's 1, 2 & 3]

Figure 5.6 Comparison between the magnitudes of the speed illusion in Experiment 1 (filled bars), Experiment 2 (open bars) and Experiment 3 (grey bars). In Experiment 1 the radial pattern had a fixed disparity throughout each trial. In Experiment 2 the disparity varied during the trial, consistent with motion through depth. In Experiment 3 disparity varied throughout a trial, and the size of dots changed consistent with the
motion-in-depth signalled by the disparity change. The pattern of results across all three experiments shows the speed illusion to be insensitive to these depth manipulations.

5.6 Discussion

It is clear from the experiments reported here that the additional depth cue manipulations had no effect on the magnitude of the speed illusion experienced when viewing a radially expanding field of random dots. The premise for this work was that if the exaggerated perception of speed was due to the interpretation of the stimulus as motion-in-depth, rather than lateral motions, then it should be possible to influence the strength of the illusion with contrary depth cues. The failure to show any effect of binocular disparity and changing size could be due to either (or both) of the following:

- the stimulus did not provide sufficiently powerful a combination of discrepant depth information to compete with the relative motion cue
or,
- the speed illusion is not related to a motion-in-depth interpretation.

Before drawing a conclusion it is worth considering how future experiments could further explore the interaction of depth cues with relative motion, and how features of the present experiments may have contributed to the lack of interaction found here. The following two sections address these points.

Other Promising Manipulations:

Vertical Disparities

One technique that could be used to contradict the relative motion cue is the manipulation of vertical disparity. Rogers & Bradshaw (1993) showed that introducing vertical disparity caused a near two-fold increase in the perceived depth of a sine grating defined by horizontal disparities. Although horizontal
disparities were fixed in this study there is no reason to suppose that a dynamic disparity stimulus would not benefit from the same manipulation.

**Occlusion**

In real life scenes the motion of a surface through depth presents to each eye differing amounts of occlusion of objects in more distant depth planes. This is referred to as da Vinci stereopsis (Anderson, 1994; Nakayama & Shimojo, 1990; ). It is possible that providing a reference pattern that not only serves to define the plane of fixation, but also provides a dynamic occlusion cue as the disparity of the dot field changes would strengthen the competing percept of motion away from the observer.

**Weakening: The Relative Strength of Cues can be Manipulated**

There is reason to believe that the weightings given to depth cues are dependent on stimulus parameters. In studies of cue interaction it has been proposed that motion cues to depth dominate disparity cues under normal viewing conditions, and attempts have been made to weaken the dominant cue. For example, Ito (1997) investigated the interplay between luminance and disparity cues to motion-in-depth as a function of the relative strength of the cues. Luminance cues were weakened in a grating defined using an RDK by a) increasing the interval between dot presentations, b) by increasing the dot displacements. Their vertically oriented, stereo-defined square grating had alternate half cycles defined to be near and far in depth using binocular disparity. When the disparity of the dots around the transition points of the grating were made to invert the stimulus was consistent with one of two interpretations: 1) the depth of the dots whose disparity had changed had moved in the z-axis, 2) the entire grating had shifted laterally. Ito found that the perception of lateral motion was favoured when ISI and dot displacement were small, but motion-in-depth was favoured when these parameters were large. These results imply that disparity cues (motion-in-depth) take second place to luminance cues (lateral motion) unless luminance cues are reduced in salience.
Johnston et al. (1994) simulated a 3-dimensional cylinder using the kinetic depth effect (KDE) combined with binocular disparity. Observers had to judge the circularity of the cylinder defined jointly by these cues. In one experiment the KDE cue to depth was heavily compromised by reducing the number of frames in the animation from more than sixteen to just two. This allowed the disparity information to be promoted. In pilot work for these experiments, using very limited dot lifetimes was effective in weakening the relative motion cue, but unfortunately had a catastrophic effect on the sense of depth due to disparity, abolishing it altogether. It is unclear how Johnston et. al., (1994) could promote stereopsis with two frame animation, so it was decided not to use this technique.

Possible Explanations for an Absence of the Hypothesised Effect

Adaptation Reduces Apparent Speed

Adaptation in vision is a well-documented phenomenon whereby the sensitivity of selected components of the visual system is reduced following prolonged exposure to a stimulus (see Mather, Verstraten & Anstis (1998) for a review). Indeed, a number of studies have made use of motion adaptation and its frequent consequence, the motion after-effect (MAE), to explore human responses to rotational and radial motions (e.g. Cavanagh & Favreau 1980; Snowden & Milne, 1996). In terms of radial motion, only expanding stimuli were presented in the experiments conducted here, and one possible criticism is that adaptation would reduce sensitivity to this pattern selectively. However, the results show no sign of adaptation effects lessening the speed illusion. In fact the biases recorded here are high in comparison with previous studies (e.g. 10-60%, as described in the introduction). In any event, a loss of sensitivity to expansion through adaptation would manifest itself as a drop in the speed illusion, an effect that has been shown here not to occur. Still, adaptation to the benchmark rotation stimulus would have the unwanted effect of exaggerating the illusion's size, which is why the normal precaution of using both directions of rotation, clockwise and counter-clockwise, was taken. The direction of rotation was
selected at random to mitigate any build-up of adaptation, which is direction-specific.

Disparity is Not a Suitable Depth Cue to Counteract Relative Motion

Disparity has been found to be a suitable cue for disambiguating relative motion signals. Imagine travelling on a train, looking through the side window. If fixating in the middle distance a shear motion is produced, with objects beyond fixation travelling 'forwards' across the retina, and those in front 'backwards'. However, the visual system is not able to ascertain the observer's motion without knowledge of the depth of the objects producing the motion signals. In vivo there are many cues to help disambiguate the information, one of which is binocular disparity. MT (Roy & Wurtz, 1990) and MST (Roy, Komatsu & Wurtz, 1992) neurons have been found in Macaque possessing two modes of directional selectivity. When crossed disparities were present, direction sensitivity was opposite to that obtained when disparities were uncrossed. The authors propose that these cells are suited to the analysis of self-motion through the environment. In view of this utility it can be appreciated that in these circumstances disparity would be a heavily weighted cue, capable of interacting strongly with motion cues. Now consider the train driver's position. His field of view contains a radially expanding optic flow. However there is no ambiguity in this information, motion of the train in the reverse direction would present a radially contracting pattern. The optic flow specifies the relative motion between observer and environment without the need to rely on disparity for arbitration. In these circumstances disparity would be a lightly weighted cue compared to RM. This suggests that disparity may not be an effective choice to bias the perception of radial optic flow. Consistent with this is the reported lack of MT neurons responsive to motion-in-depth defined by disparity (Maunsell & Van Essen, 1983), although neurons in MT (DeAngelis, Cumming & Newsome, 1998) and MST (Roy, Komatsu & Wurtz, 1992) have been shown to be sensitive to fixed disparity values. This lack of dynamic disparity sensitivity may instantiate the concept of cue veto, where information from one depth cue is paramount over
another (Bülthoff & Mallot, 1988). Psychophysically, in their study of the
detection of motion-in-depth, Sumnall & Harris (2000) concluded that disparity
signalling motion purely in the z-axis was much less powerful a cue than
combinations of disparity and lateral motion signalling motion oblique to the z-
axis. Similarly Westheimer (1990) showed that motion-in-depth based on
disparity cues alone was difficult to detect. Although Palmisano (1996) found
the addition of disparity to radial expansion augmented the sense of vection (the
perception of self-motion, Fischer & Kornmuller, 1930), this finding does not
imply that stereo depth cues would be favoured in conflicting situations.

**Eccentricity: Stimuli were Presented Peripherally**
The centre of the complex motion patterns was 2.86 degrees from the fixation
mark. Siderov & Harwerth (1995) using similar difference of Gaussian stimuli
found that (static) depth discrimination using disparity cues did not deteriorate
significantly with eccentricities up to and beyond the 2.86 degree value used
here. Of greater relevance, it has been found that the perception of motion-in-
depth is compromised to some extent by peripheral viewing (Brooks & Mather,
2000). These authors found that 4 degrees of eccentricity produced a drop in
apparent speed of a looming binary random dot pattern of some 15%. An
eccentricity of 2.86 degrees, being significantly less than 4 degrees might be
expected to reduce apparent speed by a lesser amount, and while this could have
compromised the strength of the dynamic disparity cue to a degree, probably
does not account fully for its ineffectiveness in these experiments.

**Additional Depth Cues**
Although the purpose of these experiments was to present depth cues that were
telling different stories, it must be acknowledged that extraneous depth cues
were playing a part in observers' perceptions, as discussed by Wann et. al.,
(1995). In the present work accommodation cues were not consistent with an
approaching or receding object, presumably signalling the distance to the
computer monitor. Vergence angle, too, would be in conflict, linked as it is to
both accommodation, which was fixed, and disparity, which varied in Experiments 2 and 3. The disparity sensitivity of some 25% of MST cells in monkey have been found to be affected by vergence angle (Roy, Komatsu & Wurtz, 1992) suggesting a role for this property of binocular vision. A number of experimenters have explored the importance of this issue through the use of naturalistic stimuli, rather than their simulation by computer. Some studies have found improvements in performance when experimental conditions were more true to life. For example, Buckley & Frisby (1993) found that stereo cues were more highly weighted when real objects were used instead of simulations. However, in the most recent work Bradshaw et. al., (2000) asked observers to match the dimensions of an example triangle by manipulating the locations in space of three light-emitting diodes. The authors concluded that the pattern of results was similar to those found in comparable experiments using simulated stimuli. In summary, there are theoretical and empirical reasons to suggest that non-natural viewing affects performance requiring the assimilation of multiple depth cues, though the evidence against the use of simulated stimuli is not sufficiently consistent to presume that it represents a fatal flaw in this type of work.

**Depth Cues in Everyday Scenes**

What effect is expected when depth cues are in conflict? In fact, as pointed out by Norman & Todd (1995), viewing pictures, paintings and television sets presents the viewer with numerous depth cues which are in conflict with disparity information, which (presumably) always signals the flat surface of the medium. Of particular relevance here is the experience of motion seen on a television. This frequently conflicts strongly with the disparity information about the distance and shape of the CRT display. Yet it is invariably the motion that informs the observer’s percepts, rather than the disparity. Despite this it does not follow that the current findings could have been predicted from this knowledge. In the case of a normal TV program motion cues are accompanied by many other complementary cues such as texture, shading and occlusion. With such a
weight of evidence it is not surprising that disparity, being the most discrepant

cue, would be suppressed. If it is assumed for the sake of argument that the

motion-in-depth hypothesis is true then these experiments show that relative

motion on its own is sufficient, at least under the conditions explored here, to

veto disparity cues, whether in direct conflict, or when merely signalling a

relatively infrequent deformation event, such as an expanding balloon.

5.7 Conclusions

In these experiments it has been shown that a radially expanding pattern of dots

continues to be perceived to move faster than a rotating pattern, despite the

presence of disparity and size-change depth cues indicating that the dots

occupied a fixed or receding depth plane. If the relative motion present in

expanding patterns is interpreted as motion-in-depth, then the additional depth

cues introduced here have been shown to be ineffective in influencing that

perception, implying that relative motion is a prime source of depth information.

Alternatively, the source of the speed illusion is not related to a motion-in-depth

interpretation, and remains to be explained. Geesaman & Qian (1996) suggested

that the illusion might be a function of the larger quantity of radially-sensitive

cells found in MT compared to rotationally-sensitive cells. If anything, more

cells would simply allow an increase in resolution/accuracy along the dimension

measured by such cells, rather than some distortion of reality. Although Clifford

(1999) argued on theoretical grounds in favour of motion-in-depth, empirically

the inconclusive results reported here ensure the basis of the speed illusion

remains unresolved. Regarding the second explanation (the motion-in-depth

hypothesis is incorrect), all that can be claimed at this stage is that no evidence

has been found in favour of the motion-in-depth hypothesis. However, absence

of evidence is not evidence of absence, and it remains possible that there exist

combinations of depth cues that would override RM and diminish the speed

illusion.

These results have implications for modellers of cue combination. A theme of

these models is the method by which depth cues are integrated when available in
different 'currencies'. By different currencies it is meant that some cues, for example disparity, signal relative depth while occlusion signals ordinal depth (i.e. which one is closer). After Bulthoff & Mallot's (1988) characterisation of four types of cue interaction: accumulation, veto, disambiguation and cooperation, Maloney, Landy and colleagues (e.g Landy et. al., 1995) have developed a modified weak fusion theory of cue interaction. Their model seeks to strike a balance between modular approaches to cue integration on one hand, and highly interactive approaches on the other. Modular schemes have depth estimates computed independently, then combined in an averaging process. Non-modular approaches allow full interaction between cues in simultaneous satisfaction of multiple constraints e.g. Nakayama & Shimojo (1992). These findings can be taken as further evidence that disparity is a cue with low weight where motion is along the observer's line of sight.
Chapter 6

Discrimination of Speed in Radial, Rotation and Translation Motion Mechanisms

6.1 Abstract

Knowing the speed at which you and other people or objects are moving is a critical skill for all humans in their interaction with the environment. The patterns of optic flow generated by such motions are projected on to the retina in a way that must be analysed at a global level. It is the analysis of the speed of large-scale movement that is the subject under investigation here. Mechanisms responsive to wide-field patterns of motion comprising many speeds and directions are thought to be a distinct class to local motion sensors. As such they may possess distinct tuning characteristics. Previous attempts to measure the speed tuning of such mechanisms have not distinguished between local and global mechanisms. Here three experiments are reported that address the problem of excluding local speed mechanisms from the measurement of global speed discrimination thresholds. Two techniques are employed, one to render local cues unreliable, the other to render them completely uninformative. The first technique provided estimates of speed discrimination thresholds for radial, rotation and translation (20%-70%) that were higher than many previous studies have reported. However the technique may be inappropriate for this type of work. The second technique, used with rotation only, provided lower estimates of 5%-30% that are more in keeping with the existing literature.
6.2 Introduction

The perception of simple forms of movement such as translation in the fronto-parallel plane has been explored many times. For example, studies of speed discrimination (De Bruyn & Orban, 1988; McKee, 1981; Mestre et al., 2001) in humans has shown them to be capable of determining differences of around 4% to 20% (expressed as Weber fractions), depending on the underlying speed of the reference. However, retinal stimulation produced by ego motion contains signals over large areas of the visual field, comprising a diverse range of speeds and directions. It has been postulated that special mechanisms exist in primate vision (Tanaka & Saito, 1989; Duffy & Wurtz, 1991) to process such heterogeneous signals, and previous experimental work has attempted to document the competence of such 'globally' tuned mechanisms in the detection and discrimination of speed.

Global Speed Discrimination: Translation, Rotation and Radial Motion

In one of several related experiments, Smith, Snowden & Milne (1994) found that a field of translating dots (all of the same speed) needed a velocity increase of some 10-12% to be distinguished from a reference pattern moving at 4 degrees/s. Extending earlier work on the speed discrimination of simple singular translational motions, Verghese & Stone (1995) asked observers to discriminate on the basis of speed between two animations, each comprising multiple Gabor patches. They found that speed discrimination thresholds decreased with increasing number of patches (1 patch = circa 19%; 6 patches = circa 11%, interpreted from their figure 3), a finding they attributed to the parsing of visual space into discrete entities. When the area of a single patch was increased in line with that shared across multiple patches thresholds were constant, at approximately 20% (interpreted from their figure 5a).

A particular emphasis has been placed on the ability of observers to exploit the radial pattern of optic flow produced by forward motion along the actor's line of sight. The parameters of mechanisms responsive to such radial flow have also
been thoroughly investigated. For example, much work has focussed on how accurately an observer can judge the time it will take for a simulated object to reach them, known as its 'time-to-contact' (TTC). Vincent & Regan (1997) demonstrated TTC discrimination thresholds to be around 8%. Similarly, Regan & Hamstra (1993) reported TTC discrimination thresholds for a simulated approaching object to be in the range 7-13%.

More direct attempts to measure speed discrimination thresholds in radial patterns, rather than TTC discrimination thresholds, have been made. Sekuler (1992) compared discrimination performance in rotation and 'looming' (radial expansion) patterns using random dot kinematogram (RDK) stimuli. Thresholds for both patterns were similar, ranging between 5% and 16%, depending on observer.

Clifford, Beardsley & Vaina (1999) measured speed discrimination thresholds for a relatively high reference speed (21.3 degrees/s.) in RDK's, establishing low Weber fractions of between 3% and 4% for both rotation and radial patterns. At much lower reference speeds (1, 0.5 & 0.33 degrees/s) Bex, Metha & Makous (1998) found discrimination thresholds for gratings to vary with speed, recording fractions no greater than 15%, and more typically between 5% and 10%.

Kaiser (1990) examined the ability of observers to discriminate the angular velocities of two simulated objects that rotated about axes orthogonal to the subject’s line of sight. If the three spatial dimensions are termed x-axis for horizontal, y-axis for vertical and z-axis for depth, (where the observer’s line of sight is the z-axis) the stimuli were rotated about the x and y-axes. Sometimes comparisons were made between rotations on identical axes (i.e. both objects rotating about the x-axis, or both rotating about the y-axis), sometimes between opposing axes (i.e. one rotating about the x-axis, the other rotating about the y-axis). Varying the size and shape of objects, their phase relationships and directions of rotation, the authors reported that angular velocity discrimination competence of around 12% was achieved. Matching angular velocity between objects of differing sizes showed that small objects needed to have, on average, 9% greater rotational speed to be perceived as equivalent to their larger
associates. These results, however, do not inform of the type of rotation-sensitive mechanisms that address detection of rotation about the z-axis. Nonetheless the figures reported are similar to those found by the following workers who did investigate rotation about the z-axis.

A study by Kaiser & Calderone (1991) examined angular velocity matching and discrimination in rotation about the z-axis. Using a dense field of rotating single-pixel dots they reported discrimination thresholds of between 4% and 6% when test and reference kinematograms were of equal diameter, and similar values (3%-6%) when diameters differed by a factor of two. Examining the points of subjective equality revealed no bias when kinematograms were of the same diameter, but a significant bias (17%) when diameters differed. This indicates that local velocities were influencing speed judgements, though they were clearly not completely relied upon as the size difference between rotating patterns meant that local dot speeds differed by a factor of two, implying a bias of 100% if matches were being mediated by linear motions.

Work on rotational speed tuning (speed matching and speed discrimination) centred on the z-axis has also been reported in three papers by Werkhoven & Koenderink (1990, 1991, 1993) using random dot stimuli. Their first study examined speed discrimination between rotating annuli, showing that lowest difference thresholds of 8-10% were achievable, within the range for local judgements (De Bruyn & Orban, 1988; and McKee, 1981). In their second paper (Werkhoven & Koenderink, 1991), a series of four experiments examined the dependency of speed discrimination on several factors: temporal factors of frame duration and trial duration; textural factors, i.e. number of dots and dot density; spatial scale, where both comparison stimuli were presented at a range of viewing distances to manipulate their retinal sizes (test and reference always the same size); and finally, the dependence of discrimination threshold on speed was explored to ascertain whether Weber's law held for rotational motion in the way it does for linear motion. The authors found that speed discrimination performance in all experiments was similar to that previously found for linear motion. They concluded that observers were not able to make use of angular
velocity to judge difference thresholds, and were therefore basing their decisions on the local retinal velocities, averaging these where multiple speeds were present in the stimulus.

In their later study, Werkhoven & Koenderink (1993) further explored the effect of spatial scale by assessing how perceptual matches and speed discrimination thresholds were influenced by differences in the radii of test and reference patterns. This was done for a range of speeds, using rotating random dot stimuli. Discrimination thresholds were found to vary with the degree of mismatch between the radii of comparison stimuli. Minimum threshold was obtained when sizes were equal (7%), increasing to more than 20% when radii differed by a factor of two. Again, the authors concluded that fundamentally the local linear speed of dots was being used in preference to the angular speed of rotation.

Targeting global mechanisms in an attempt to measure speed discrimination thresholds encounters a major problem implied by the ‘multiple stage’ theory of motion sensing currently in favour (Adelson & Movshon, 1982; Welch, 1989). This model addresses the problem of how an initially piecemeal analysis of the visual scene can be used to appreciate events in the world whose consequences span wide fields of view. One simple formulation of the idea has a battery of spatially localised motion detectors at the first stage, with small receptive fields tuned for specific directions of motion, feeding into a second stage mechanism that integrates across the inputs from the first stage. If the directional properties of first stage motion detectors were carefully arranged spatially to match the distribution of speeds and directions of, say, a rotation pattern, then the second stage mechanism would inherit a large receptive field tuned for just such a stereotypical rotation event. Such models are described as templates, since they simply respond when there is a match between them and the external stimulus (e.g. Perrone & Stone, 1998). However, as first stage mechanisms have motion-sensing competence of their own, it is difficult to be sure that attempted measurements of second stage global detectors are not in fact tapping the first stage. Sekuler (1992) concluded that speed sensitivity to complex forms of
motion was indeed mediated by simple pooling over local mechanisms, as illustrated in a quantitative modelling exercise they carried out. Werkhoven & Koenderink (1993) came to similar conclusions, that fundamentally local linear speeds of dots were being used in preference to the angular speed of rotation of the patterns. The speed discrimination results (3%-4%) of Clifford, Beardsley & Vaina (1999) and Kaiser & Calderone (1991) (3%-6%) are perhaps the most similar to previous local motion discrimination figures (e.g. 5%, McKee, 1981).

It was decided to address the problem of precluding the use of local motion detectors in the assessment of speed discrimination performance in global patterns of motion. Two methods have been employed. In Experiment 1 a signal-in-noise paradigm is used, in combination with ‘speed balancing’ to discourage reliance on local motion signals. A signal-in-noise task involves recording the performance of an observer as a function of the ratio of signal to noise information contained in the stimulus. The ratio of signal in the patterns used here was 50%, considerably above normal motion coherence threshold of circa 5% (Scase, Braddick & Raymond, 1996). In Experiments 2 and 3 the usual connection between local and global speeds is broken to achieve the same objective. These techniques are described in the relevant experimental sections. Experiment 1 explores discrimination thresholds in global patterns of translation, rotation and radial motion, whereas Experiments 2 and 3 are limited to the investigation of rotation.

6.3 Experiment 1. Speed Discrimination Thresholds for Global Motion: Translation, Rotation and Radial Motion

As outlined in the Introduction, the problem for previous research has been the possibility that observers were able to make use of local linear dot speeds to discriminate speed differences between stimuli. To address this problem a method was chosen to oblige subjects to abandon local speed cues in favour of global speed cues. Observers were presented with two animations containing global motion (e.g. rotation) and were required to identify which one appeared faster (see figure 6.1). When two RDK’s each contain a rotation, the faster one
can be identified, in theory at least, by identifying the animation containing either the highest local speed(s), or the highest average speed, regardless of the global structure.

Figure 6.1 Example of stimuli used in Experiment 1, rotation condition. Two Random Dot Kinematograms (RDK’s), were presented, each containing a global pattern of motion. The observer’s task was to choose which pattern rotated faster. Two other global motion patterns were used, radial (expansion and contraction) and translation. Both animations were carefully matched for average speed, despite the differing rates of rotation by, ‘speed balancing’. See text and figure 6.2 for details.
Figure 6.2 ‘Speed balancing’ in the signal-in-noise stimulus. When two animations containing global motion are presented in the speed discrimination task of Experiment 1, the signal speed of one pattern, the ‘Reference’ is fixed (Panel B). The signal speed in the ‘Test’ pattern is manipulated (Panel A, double headed arrow) to generate a psychometric function. This would normally result in a difference between the average speeds of the test and reference patterns that could be used to discriminate between them. To remove this we manipulated the speed of the ‘noise’ dots in the reference pattern (Panel B, double headed arrows) in concert with the signal speed in the test pattern. This manipulation ensured that both animations shared the same average and peak speeds at all times. Speed profiles are triangular in shape reflecting how dot density varies with eccentricity, i.e. there are more high-speed dots at the outer
circumference of a rotating disc than there are near its centre. More dots appear in the periphery, and they travel at greater speed.

To prevent either of these strategies from being effective a ‘speed balancing’ procedure was adopted that equated the animations for speed at all times. Both animations depicting, say, rotation were immersed in a population of noise dots. Noise dots moved with a range of speeds similar to the rotating dots, however their directions were assigned at random.

Now, when the speed of the rotating dots in the ‘Test’ animation is varied, the subsequent elevation of peak and average dot speeds can duplicated in the alternative animation by varying the speed of the noise dots in the ‘Reference’ window (see figure 6.2). As there were the same number of signal dots as noise dots in each animation the average dot speed between the animations was equated at all times, and observers were deprived of these cues to aid them.

6.3.1 Methods

Stimuli

Dot specifications. Each Difference of Gaussian dot subtended approximately 0.753 degrees of visual angle.

In total 150 dots appeared in each animation, 75 signal dots and 75 noise dots.

To create rigid forms of rotation and radial motion the speed of each dot in these patterns was scaled linearly to its distance from the centre of the flow field, yielding biologically relevant speed gradients. Two conditions of speed were explored. The maximum speed present in the reference pattern seen at the outside extremity of the animation window was 5.23 degrees/s (high-speed condition), and 3.14 degrees/s (low-speed condition). The median of the speed range of the reference pattern in all three forms of motion in the fast condition was therefore 3.14 degrees/s, and 1.57 degrees/s for the slow condition. As
translation contained no speed gradients only one speed appeared in the pattern, which was equal to the median speed in the complex patterns.

**Procedure**

Observers viewed two RDK patterns in a two (spatial) alternative, 'forced choice' paradigm. One pattern was of fixed speed of global motion ('reference'), the other of variable speed ('test') (see figure 6.1).

Three forms of global motion were explored in this experiment: rotation, radial and translational motion. The following description for rotational stimuli may be extended to include the other patterns. Each pattern comprised two populations of dots. One population ('signal') rotated around the centre of the animation window, the other ('noise') had randomly assigned directional components. The subjects' task was to nominate which animation appeared to rotate faster.

The method of constant stimuli was used to manipulate the speed of the test pattern.

A block of trials comprised sixty observations, with at least four blocks of data being collected, resulting in a minimum of 240 trials per psychometric function generated. No feedback was given regarding the accuracy of the observers’ responses.

Subjects were three experienced psychophysics observers (AS, PB and RW), the author being RW.

Different forms of global motion were presented in separate sessions.

**6.3.2 Results**

Difference thresholds were estimated from the cumulative Gaussian best fitting the observers' psychometric functions. Examples of typical psychometric functions with the best fit and 95% confidence interval fits are shown at figure 6.3 for all three forms of global motion presented in the high-speed condition (median speed =3.14 degrees/s).
Figure 6.3 Examples of psychometric functions relating the probability of seeing the test pattern as faster (y-axis) to the median speed of the test pattern (x-axis). Panels A & B show data for the Radial condition. (NB Observer AS was unable to perform this task). Panels C, D & E show data for the Rotation condition, and panels F, G & H show the
Translation data. Blue diamonds signify the empirical data points, the best-fitting cumulative Gaussian (solid purple line) is flanked by curves delimiting the 95% confidence interval.

Speed discrimination thresholds for both observers are plotted as a function of global motion type in the histograms of figure 6.4. Panel A shows results for the slow speed condition (1.57 degrees/s., panel B for the high-speed condition (3.14 degrees/s). (N.B. observer AS was unable to perform the task for radial motion in the high-speed condition). Thresholds are expressed as a percentage of the median speed of the reference pattern, indicating that the test animation needed a local median dot speed increase of between 20% and 50% to be reliably perceived as faster. It can be seen that the difference in the height of the bars indicates substantial individual differences, though the general trend is similar for both. More striking is the magnitude of the thresholds. For observer AS these range between 50% and 86%, and for PB 41% and 50%, both very large compared to thresholds previously reported (e.g. 5-16%). Even for RW lower values of between 16% and 35% are still a good deal greater than found elsewhere. Such large Weber fractions appear to justify the attempt to measure the speed competence of global mechanisms. Large Weber fractions for speed discrimination have been reported previously, using quite different methods. Banton & Bertenthal (1992, conference abstract) asked subjects to discriminate the orientation of a rectangle made up of spinning line segments, embedded in a field of similar rotating lines. When the speed of the lines within the rectangle was sufficiently greater than that of the surround the rectangle stood out clearly. This method showed that angular speed differences between populations of spinning elements needed to be 40% or more to achieve judgements of the orientation of the rectangle. Snowden & Braddick (1991) presented participants with two RDK patterns. In one all dots had a single velocity component, in the other the dots took on one of two velocities interleaved in time. The task was to identify which animation contained the velocity oscillation. Varying the difference between velocities in the oscillating animation, Snowden & Braddick
(1991) found that discrimination thresholds of between 25% and 40% were obtained, with higher thresholds accruing at higher oscillation rates. Evidence from single cell physiological studies in primates suggests that speed tuning in MST, the site most strongly associated with optic flow processing, is not narrow. For example both Tanaka & Sato (1989) and Duffy & Wurtz (1991) found that individual neurons were relatively insensitive to speed changes of the order of tens of degrees/s, this being so over a wide range of stimulus speeds (e.g. 10-80 degrees/s). Duffy & Wurtz (1991) speculated that this was consistent with the nature of realistic optic flow, where a broad range of speeds is simultaneously present in typical circumstances, i.e. a speed gradient. This, of course, does not exclude the possibility that a population code (see Pouget, Dayan & Zemel, 2000 for a review of population coding) among broadly tuned cells could not resolve speed to fine degrees, yielding superior performance at a broader level of analysis (Duffy & Wurtz, 1991), or that later visual areas to which MST projects might not do so (e.g. Anderson & Siegel, 1999; Phinney & Siegel, 2000).
Figure 6.4 Histogram summaries of the speed discrimination thresholds obtained for each observer. Panel A shows thresholds from the low-speed condition (median speed = 1.57 degrees/s.); Panel B shows thresholds from the high-speed condition (median speed = 3.14 degrees/s.).
speed = 1.57 degrees/s), panel B shows the same information from the high-speed condition (median speed = 3.14 degrees/s). Error bars are 95% confidence intervals.

However, it is possible that these large Weber fractions do not reflect insensitivity to speed at the global level. Previously Watamaniuk & Duchon (1992) have established that when multiple speeds are present the visual system averages speed information and bases discrimination judgments on the mean speed of all dots in the pattern. Their data show this to be true where the range of speeds in the pattern was 2.2-8.5 degrees/s. On the assumption that such a strategy will eventually be discarded, as it must in order to perceive transparency, then the measurements made here may reflect the change from an averaging strategy to a segmenting strategy (see Treue, Hol & Rauber (2000) for a population based account of signal segmentation in the direction domain). Masson, Mestre & Stone (1999) documented speed segmentation thresholds for two populations of unidirectional dots travelling at different speeds. Dependency on display duration was found, with thresholds of circa 5%-20% being registered at their greatest tested duration of 500mS for dot speeds comparable to those presented here (c.f. duration here of 1000mS). Direct comparison with their results is not possible, however, as the dots used here were multi-directional. Nevertheless, it is plausible that under conditions where prior segmentation is required, discrimination thresholds might be expected to be high.

Calculating the greatest range of speeds seen in the test pattern (by PB, who required the largest difference between reference and test to generate a complete psychometric function) shows that speeds in the range 3-4.97 degrees/s. were present. This range is within that found to result in averaging by Watamaniuk & Duchon (1992) in their multi-directional random walk RDKs. However, unlike their study the RDK’s presented in this experiment did contain a population of directionally coherent dots, and so their data do not predict over what range of speeds averaging might be expected to take place. Conceivably then, the stimulus design used here to prevent reliance on local cues has also compromised the use of the global structure. This despite the fact that the signal
ratio (50%) was well above the level required to detect optic flow components in noise (Scase, Braddick & Raymond, 1996). Still, detecting such optic flow components lies in the ability to segment on the basis of direction within the same interval. It does not follow that discriminating between two intervals on the basis of speed is then made as straightforward as it would be with no noise components present. It appears that the presence of noise dots continues to influence the subject's responses, therefore leaving a question mark over the suitability of this approach to measure global speed.

It was decided that this possibility demanded a fresh approach.

6.4 Experiment 2
Speed discrimination thresholds found in Experiment 1 were high when compared with those previously reported in the literature. Therefore a further experiment was designed, aimed at assessing speed discrimination thresholds using a different methodology. The primary objective remained the same i.e. to ensure global mechanisms were being used to mediate the observers' judgements. To this end it was decided to ask subjects to compare the speeds of two RDK's depicting rotating annuli (see figure 6.5). Comparisons between rotating annuli RDK's of the same radius confounds judgements based on local dot speeds with those mediated by rotational speed, making it impossible to attribute the source of performance with confidence. For example, to have two patterns of identical radius rotating with an angular speed differential of 10% requires a local dot speed differential of 10% also. To address this problem the connection between local and global speeds was broken by comparing between rotating annuli of differing radii. If one annulus were twice the radius of its confederate, but both contained dots travelling at the same local speed, the smallest of the pair would rotate at twice the angular velocity of the other. If the visual system makes use of special mechanisms sensitive to angular velocity then the smaller will indeed appear to travel at twice the speed of the larger. If not, the visual system may make use of local dot speeds, yielding an apparent match between the patterns. In this way it was hoped to be able to establish
whether speed discrimination thresholds were mediated by rotation-sensitive mechanisms. To further promote the use of global mechanisms very short lifetime dots (life=2) were used, depriving local mechanisms of temporally and spatially elongated trajectories over which to operate.

Figure 6.5 Example of stimuli used in Experiment 2. A pair of concentric annuli equated for stimulus area, were presented. Each comprised a field of 200 randomly positioned Difference of Gaussian dots. Both annuli shared the same direction of rotation, clockwise or anti-clockwise, chosen at random on each trial. The observers judged which annulus appeared faster. Three conditions of annulus diameter were used: Panel A ‘Large-Medium’; Panel B ‘Large Small’; Panel C ‘Medium-Small’. The larger
diameter annulus was always the fixed-speed reference pattern. See Methods for dimension details.

As it was uncertain at the outset whether subjects would use global or local cues, it was necessary to establish psychometric functions for the points of subjective equality (PSE) between annuli of different radii. If the PSE's showed that observers were using angular velocity to match the patterns then estimating discrimination thresholds from the functions obtained would be justified. Previous work exploiting similar ideas from Kaiser & Calderone (1991) employed rotating discs of unequal radii, a stimulus design in which a large range of local speeds exists as a dot's speed depends on its eccentricity. They found a tendency to use angular speeds, but with a bias towards local speeds. The annuli used here were of a narrow profile to restrict the range of local speeds and hence simplify the interpretation of the results. Werkhoven & Koenderink (1993) used rotating annuli similar to those used here but with long lifetimes, but have been criticised by Barazza & Grzywacz (2001) on the grounds that their stimuli were insufficiently dense to promote reliance on angular velocity information.

Although a similar case can be made for dissociating local and global speeds in radial patterns by using display windows of differing diameter, a theoretical problem argued against this technique. The 3d velocity (motion-in-depth) interpretation of a radial pattern is ambiguous without a depth to scale the distance of the pattern elements. To elaborate, any given radial pattern could be consistent either with a far object approaching quickly, or a near object approaching slowly. Although this is true of any two radial patterns regardless of diameter, most experiments make comparisons between similar designs of stimuli, and the same depth estimate could be expected for each. It is not certain that the visual system would treat unequally sized radiations in the same way. As for translation, there is no obvious way to dissociate local and global speeds.
6.4.1 Methods
Observers viewed two rotating RDK annuli in a two (spatial) alternative, 'forced choice' paradigm (see figure 6.5). Each pattern comprised a population of dots rotating around the centre of the annulus, one of fixed speed of rotation ('reference'), the other of variable speed ('test'). The subject’s task was to nominate which animation appeared faster. A Quest algorithm (Watson & Pelli, 1983) was used to manipulate the speed of the test pattern. No feedback was given as the task was of a subjective nature.

Annulus Sizes, Locations and Speeds
Three sizes of annulus were used, as specified below. For ease they will be referred to as large, medium and small.

‘Large’. (Used as reference only). The outer dimension of the annulus subtended 6.67 degrees of visual angle at the viewing distance of 57cm. The inner dimension was 5.93 degrees.
‘Medium’. (Used as test and reference). Outer annulus dimension 5.337 degrees and inner 4.37 degrees.
‘Small’. (Used as test only). Outer annulus dimension 3.8 degrees and inner 2.23 degrees.
In all cases the area of the annuli was 30 degrees. As 200 dots were displayed in each pattern, dot density was 6.667 dots/degree in each case.

Four conditions of speed were explored, which were randomly interleaved within a block of trials. In all experiments the larger diameter annulus (of the chosen pair) served as the reference. The nominal dot speeds used in the program were those that would appear at the very limit of the circular window in which the animation was displayed. The speed of an individual dot was scaled to its eccentricity to produce a rigid rotation within each annulus. Therefore speed
gradients were present, and for this reason the median speed present in each reference annulus is quoted, for each of the four speed conditions:
Large: 4.7, 7.1, 9.4, 11.8 degrees/s.
Medium: 3.6, 5.4, 7.25, 9.1 degrees/s.
Annuli were displayed concentrically. Both were placed in the centre of the screen and were presented simultaneously, though this arrangement precludes randomisation of the spatial location containing each pattern. The directions of rotation were the same for both annuli, the direction being chosen at random on each trial. A black fixation cross was presented measuring 1.34*1.34 degrees, located at the centre of the annuli.
Other methodological details were as for the previous experiment.

Procedure
The observer's task was to nominate which annulus appeared to rotate faster. If the subject perceived the inner annulus to be faster button '1' was pressed, if the outer annulus appeared faster then button '3' was pressed. A block of trials comprised twenty observations, with at least four blocks of data being collected, resulting in a minimum of 80 trials per psychometric function generated. The adaptive nature of the Quest algorithm meant that the number of observations per data point was not a constant quantity.

6.4.2 Results
Figure 6.6 shows for each observer how the points of subjective equality (blue diamonds), on the y-axis, vary with the speed of the reference pattern, on the x-axis. A linear regression (solid black line) has been fitted to the data, with its equation and $R^2$ values displayed on each graph. In all cases the regression line provides a good fit to the data, $R^2$ never being less than 0.94. Also shown on the plots are predictions for the PSE's if local dot speed were used to do the task (solid yellow line), and the predictions were angular speed to be used (solid purple line). It can be seen that for two observers (PB and AS) the data falls part
way between these predictions, whereas the author's data (RW) tends to agree with the prediction for angular speed matching. It would seem that the ability to use angular velocity in this type of task is not mandatory, and may require training of observers to consult the most advantageous mechanisms. Interestingly, despite evidence for rotational speed matching in some conditions, there were never any instances where matching was clearly and unambiguously based on linear speed. This indicates that linear speed matching is not mandatory either. The fact that matching tends to take place between angular and linear speeds indicates that both speeds are available to observers, but that they have some difficulty selecting between them, effecting a compromise in their decision.

While this experiment was being conducted evidence in a similar task exploring speed matching (but not speed discrimination) was reported from another laboratory. Barazza & Grzywacz (2001) found that all four of their subjects were able to match rotating RDK annuli of different radii on the basis of angular speed. Therefore it was decided to estimate speed discrimination thresholds from the data sets where observers were performing the task using angular velocity, i.e. where points of subjective equality indicated use of angular velocity rather than local element speed. In common with other investigators the point where the test pattern was judged faster on 75% of trials was chosen (e.g. Clifford, Beardsley & Vaina, 1999), though it is not uncommon for other performance criteria to be used (e.g. 71%: Sekuler, 1992; 82%: Verghese & Stone, 1995; 84%: Werkhoven & Koenderink, 1991). All of observer RW's data were used, but none of AS's data or PB's data were suitable. Table 6.1 shows speed discrimination thresholds for RW as a function of the reference speed and the sizes of the annuli. The range of thresholds was 4.5% to 13.4%. The means shown in the outer row and column indicate that there is a tendency for lowest thresholds (mean=6%) to be associated with 9.88 degrees/s reference speed, and for the Large-small configuration to obtain the highest (mean=11%). These values are very much in accordance with previous estimates of difference thresholds in the literature, and somewhat lower than the values found in
Experiment 1 (lowest thresholds for RW= circa 20\%..30\% for in the rotation condition).
PSE (pb concentric) Large-medium

\[ y = 0.7968x + 1.1218 \]

\[ R^2 = 0.9913 \]

PSE (pb concentric) Large-small

\[ y = 0.776x - 0.0398 \]

\[ R^2 = 0.9647 \]

PSE (pb concentric) Medium-small

\[ y = 0.8131x - 0.006 \]

\[ R^2 = 0.9625 \]
PSE (rw concentric) Large-medium

- Match speed
- Angular speed prediction
- Linear speed prediction
- Regression line

\[ y = 0.5703x + 1.4218 \]
\[ R^2 = 0.9906 \]

Reference speed (deg/s)

PSE (rw concentric) Large-small

- Match speed
- Angular speed prediction
- Linear speed prediction
- Regression line

\[ y = 0.3397x + 1.2668 \]
\[ R^2 = 0.9609 \]

Reference speed (deg/s)

PSE (rw concentric) Medium-small

- Match speed
- Angular speed prediction
- Linear speed prediction
- Regression line

\[ y = 0.6093x + 0.3406 \]
\[ R^2 = 0.9866 \]

Reference speed (deg/s)
Figure 6.6 Points of subjective equality (PSE) between concentric rotating annuli of unequal diameter (Experiment 2). For each observer the curves show how the PSE’s (y-
axis) vary with the mean speed of the reference pattern (x-axis). The reference was always the larger of the pair. Data (blue diamonds) are fitted with a linear regression, whose equation and $R^2$ value are reproduced on the graphs. The solid yellow line is the performance predicted if the subjects were matching pattern speeds using local linear dot speeds. The solid purple line shows the predicted match if angular velocity were being used.

<table>
<thead>
<tr>
<th>Reference speed (deg/s)</th>
<th>4.94</th>
<th>7.41</th>
<th>9.88</th>
<th>12.35</th>
<th>mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large-medium (RW)</td>
<td>7.8%</td>
<td>8.8%</td>
<td>6.5%</td>
<td>6.5%</td>
<td>7%</td>
</tr>
<tr>
<td>Large-small (RW)</td>
<td>9.0%</td>
<td>13.4%</td>
<td>7.6%</td>
<td>12.4%</td>
<td>11%</td>
</tr>
<tr>
<td>Medium-small (RW)</td>
<td>11.8%</td>
<td>7.4%</td>
<td>4.4%</td>
<td>4.5%</td>
<td>7%</td>
</tr>
<tr>
<td>mean</td>
<td>10%</td>
<td>10%</td>
<td>6%</td>
<td>8%</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.1 Speed discrimination thresholds obtained with concentric rotating annuli. Thresholds are expressed as the percentage speed increase required by the test pattern over the reference pattern for the test to be judged faster on 75% of trials. This speed was calculated by taking the inverse of the best-fitting cumulative Gaussian function to the subject's psychometric function. Thresholds were only calculated where the subject's performance on the task was consistent with the use of angular velocity, i.e. when the point of subjective equality across a range of speed conditions qualitatively matched the point of objective equality predicted by an angular velocity matching strategy.

6.5 Experiment 3

There exists a growing literature on the centre surround properties of MT motion cells, suggesting that the response of some mechanisms is modulated by the presence of motion in their surround. Although this has been interpreted as serving a role in scene segmentation, as motion boundaries occur at depth changes and object borders (Allman et. al., 1985a,b; Lamme, 1995; Eifuku & Wurtz, 1998; Eifuku & Wurtz, 1999) it is possible that speed perception is influenced by the concentric arrangement employed in the previous experiment. Also, Werkhoven & Koenderink (1990) examined the effect of a concentric
interference annulus on speed discrimination in rotating patterns, finding that discrimination thresholds were elevated by its presence. In those experiments test and reference annuli were displayed in two temporal intervals, each with its accompanying interference ring located concentrically. The detrimental effect on speed discrimination performance was attributed to the size of local motion receptive field sizes spanning both patterns. This interpretation is in line with their wider conclusion that speed discrimination is based on local linear signals, rather than rotational speed. For these reasons the spatial arrangement of the annuli was altered to ensure that the results from the previous experiment were not peculiar to the stimulus organisation.

6.5.1 Methods
In the third experiment the only change made was to the spatial organisation of the rotating annuli. Instead of being placed concentrically they were arranged to be side by side (see figure 6.7). The dimensions of the patterns were as for Experiment 2.
The centre point of each annulus was placed in the centre of the screen in the vertical dimension, and displaced 7 degrees from the screen centre in the horizontal dimension. The spatial locations (left/right) of the animations were assigned at random on each trial. A black cross measuring 1.34*1.34 degrees was placed in the centre of the screen, which the subjects were required to fixate during trials. Observers used button ‘1’ on the keyboard to signify that the left location looked faster, and button ‘3’ the right location as faster.
Figure 6.7 Example of stimuli used in Experiment 3. A pair of adjacent annuli that were equated for stimulus area were presented. Each comprised a field of 200 randomly positioned Difference of Gaussian dots. Both annuli shared the same direction of rotation, clockwise or anti-clockwise, chosen at random on each trial. The observers judged which pattern appeared faster. Three conditions of annulus diameter were used: Panel A ‘Large-Medium’; Panel B ‘Large-Small’; Panel C ‘Medium-Small’. The larger diameter annulus was always the fixed-speed reference pattern. See Methods for dimension details.

### 6.5.2 Results

Figure 6.8 shows for each observer how the points of subjective equality vary with the speed of the reference pattern. Predictions are again plotted for PSE’s mediated by local dot speed (solid yellow line), and for angular speed (solid purple line). The results with the non-concentric configuration are qualitatively similar to those found with concentric annuli, with mixed evidence for the use of rotational (angular) velocity in making speed matches. These results are similar to those from Experiment 2 therefore it appears that the data from the previous experiment were not compromised by the spatial arrangement of the patterns. Observer RW shows greatest consistency, with the data falling close to the angular velocity predictions in all three conditions. Both PB and AS show speed matching based on angular velocity in the ‘Large-medium’ condition only. For the remaining conditions the speed matching of these two observers falls part way between the angular and linear predictions, again reminiscent of the original findings of Werkhoven & Koenderink (1993).
PSE (pb non-concentric) Large-medium

\[ y = 0.9135x - 0.7533 \]
\[ R^2 = 0.9727 \]

PSE (pb non-concentric) Large-small

\[ y = 0.7397x - 0.556 \]
\[ R^2 = 0.9569 \]

PSE (pb non-concentric) Medium-small

\[ y = 0.9442x - 0.7699 \]
\[ R^2 = 0.989 \]
PSE (rw non-concentric) Large-medium

- Match speed
- Angular speed prediction
- Linear speed prediction
- Regression line

$y = 0.5786x + 1.3582$
$R^2 = 0.9784$

PSE (rw non-concentric) Large-small

- Match speed
- Angular speed prediction
- Linear speed prediction
- Regression line

$y = 0.5163x + 0.0223$
$R^2 = 0.97$

PSE (rw non-concentric) Medium-small

- Match speed
- Angular speed prediction
- Linear speed prediction
- Regression line

$y = 0.6395x + 0.3922$
$R^2 = 0.9761$
PSE (as non-concentric) Large-medium

- Match speed
- Angular speed prediction
- Linear speed prediction
- Regression line

\[ y = 0.8441x - 0.3517 \]
\[ R^2 = 0.9911 \]

PSE (as non-concentric) Large-small

- Match speed
- Angular speed prediction
- Linear speed prediction
- Regression line

\[ y = 0.6707x - 0.3595 \]
\[ R^2 = 0.9827 \]

PSE (as non-concentric) Medium-small

- Match speed
- Angular speed prediction
- Linear speed prediction
- Regression line

\[ y = 0.8344x - 0.5779 \]
\[ R^2 = 0.98 \]
Figure 6.8 Points of subjective equality (PSE) between non-concentric rotating annuli of unequal diameter (Experiment 3). For each observer the curves show how the PSE’s (y-axis) vary with the mean speed of the reference pattern (x-axis). The reference was always the larger of the pair. Data (blue diamonds) are fitted with a linear regression, whose equation and $R^2$ value are reproduced on the graphs. The solid yellow line is the performance predicted if the subjects were matching pattern speeds using local linear dot speeds. The solid purple line shows the predicted match if angular velocity were being used.

As in Experiment 2, it was decided to extract discrimination thresholds from the data of observers who were unambiguously making use of global cues in the task. All of observer RW’s data and the ‘Large-medium’ data of AS and PB were used. Table 6.2 shows that speed discrimination thresholds range between 4.6% (RW) and 28.9% (AS). This is a greater spread of values than found in Experiment 2. Clearly observers PB and AS show the highest values, tending to be worst at low speeds. The histogram of figure 6.9 shows these tendencies graphically.

Examining the data of RW shows that the means are comparable to those of Experiment 2, though a little higher on average. The same trend for highest thresholds is seen with the ‘Large-small’ configuration, and the lowest are associated with the higher reference speed conditions. The mean discrimination thresholds as a function of stimulus speed (bottom row, 28.9% to 13.9%) show a marked tendency to decrease with increasing speed, pointing to a failure of Weber’s law. This is not a unique finding as Masson, Mestre & Stone (1999) observed a similar phenomenon in their study of discrimination thresholds in transparent motion using random dot stimuli at speeds of up to 16 degrees/s. Heidenreich & Turano (1996) also documented discrimination thresholds that fell systematically with rising speed using grating stimuli up to 8 degrees/s. The data in the present experiment are more pronounced than in Experiment 2, where mean discrimination thresholds fell from 10% at the lowest speed, through 6%, to 8% at the highest (see table 6.1). Nevertheless, a trend does appear to be
emerging for rotation discrimination thresholds to diminish with increasing speed, as they do in studies employing linear motion (Heidenreich & Turano 1996; Masson, Mestre & Stone 1999).

<table>
<thead>
<tr>
<th>Reference speed (deg/s)</th>
<th>4.94</th>
<th>7.41</th>
<th>9.88</th>
<th>12.35</th>
<th>mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large-medium (RW)</td>
<td>15.7%</td>
<td>4.8%</td>
<td>5.6%</td>
<td>10.7%</td>
<td>9.2%</td>
</tr>
<tr>
<td>Large-small (RW)</td>
<td>16.6%</td>
<td>13.3%</td>
<td>8.6%</td>
<td>8.4%</td>
<td>11.73%</td>
</tr>
<tr>
<td>Medium-small (RW)</td>
<td>15.7%</td>
<td>10.7%</td>
<td>6.1%</td>
<td>4.6%</td>
<td>9.28%</td>
</tr>
<tr>
<td>Large-medium (PB)</td>
<td>24.7%</td>
<td>10.75%</td>
<td>18%</td>
<td>16.9%</td>
<td>17.59%</td>
</tr>
<tr>
<td>Large-medium (AS)</td>
<td>28.9%</td>
<td>23%</td>
<td>15.1%</td>
<td>13.9%</td>
<td>13.9%</td>
</tr>
<tr>
<td>mean</td>
<td>20.32%</td>
<td>12.51%</td>
<td>10.68%</td>
<td>10.9%</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.2 Speed discrimination thresholds obtained with non-concentric rotating annuli. Thresholds are expressed as the percentage speed increase required by the test pattern over the reference pattern for the test to be judged faster on 75% of trials. This speed was calculated by taking the inverse of the best-fitting cumulative Gaussian function to the subject's psychometric function. Thresholds were only calculated where the subject's performance on the task was consistent with the use of angular velocity, i.e. when the point of subjective equality across a range of speed conditions qualitatively matched the point of objective equality predicted by an angular velocity matching strategy.
Figure 6.9 Summary histogram of speed discrimination thresholds as a function of observer and viewing configuration that resulted in matching using angular velocity.

6.6 Discussion

Three experiments have been performed in an attempt to measure the speed discrimination competence of global motion mechanisms. Two different methodologies were employed to preclude the use of local speeds in performing the tasks. Experiment 1 used a signal-in-noise paradigm, forcing observers to integrate over many individually uninformative speed signals, segmenting the global motion from directionally noisy motion signals. Experiments 2 & 3 disrupted the usual identity between local and global speeds, allowing the source of information used in the task to be determined. Experiment 1 dealt with three classes of global motion, namely translation, rotation and radial motion, whereas Experiments 2 & 3 examined only rotation, as the method employed was best suited to that pattern.
The results from the two different methodologies were inconsistent. Experiment 1 found discrimination thresholds as high as 80%, and as low as 14%, a consequence mostly of individual differences. For rotation only the values were between 19% and 60%. Experiments 2 & 3 yielded lower thresholds for rotation of between 4.5% and 28.9%, although, again, individual differences accounted for much of the spread. The data from Experiment 1 has limited support from the existing literature. The form discrimination task of Banton & Bertenthal (1992) did show high thresholds for isolating the orientation of a rectangular arrangement of spinning lines, but it is not obvious how this measure relates to the operational definition of threshold for speed discrimination (75% ‘faster’). While electrophysiological findings of global pattern cells broadly tuned for speed (Tanaka & Sato, 1989; Duffy & Wurtz, 1991; Orban et. al. 1995) lends support for high discrimination thresholds, the possibility that population coding mediates psychophysical performance leads to some doubt that the findings of Experiment 1 are robust. These findings do not represent a threat to the use of the signal-in-noise paradigm per se, as long as it is used to document relative changes in thresholds across conditions. However its use for absolute measures of performance such as attempted here may be questioned.

Experiments 2 & 3 returned discrimination thresholds similar to that reported previously, while correcting for some of the compromises inherent in those studies. However, it is noted that individual differences have a major role to play in the estimates obtained, and the ability of the more practiced observer in this study to selectively use angular velocity implies that flexibility is being exercised in the monitoring of appropriate stimulus characteristics. The PSE’s across a range of annuli diameters and speeds provides confirmation that global speed was being used, supporting the validity of the discrimination thresholds that were subsequently estimated.

In conclusion, it was found that discrimination thresholds for global motion mechanisms are likely to be, as reported previously, in the region of 5% to 20%, at least under the conditions employed here. There is also some evidence that Weber’s law does not hold, as difference thresholds diminish with increasing
mean speed, a result echoed by Masson et. al. (1999) and by Heidenreich & Turano (1996).
Chapter 7
Do Relative Motion and Binocular Disparity Interact in the Detection of Shear Motion?

7.1 Abstract
There are many circumstances in which retinal motion signals are useful in calculating the direction of self-motion. However such information can be ambiguous, requiring additional cues to resolve the meaning contained in the optic flow signature. The following experiments investigated how binocular disparity depth cues influence the detection of shear (spatially separated, opposing linear motions). In contrast to psychophysical (Hibbard, Bradshaw & Hibbard, 1999; Snowden, & Rossiter, 1999) and electro-physiological data (Roy, Komatsu & Wurtz, 1992; Roy & Wurtz, 1990), no interaction was found between these attributes under the conditions explored.
7.2 Introduction

It has been known for some time that the perception of a slanted surface is indicated by a velocity gradient that varies with the degree of slant (Braunstein, 1968; Harris, Freeman & Hughes, 1992; Meese, Harris & Freeman, 1995; Freeman & Fowler, 2000). Of course, motion is only one of many cues that contribute to the perception of depth, binocular disparity being another that has been shown to produce a sensation of surface slant in humans (Howard & Kaneko, 1994) and in primates (Taira et al., 2000). Unsurprisingly, such useful depth cues have been found to interact in tasks such as 3-D shape judgements (e.g.; Bradshaw & Rogers, 1996), and there is physiological evidence that cells in primate cortical areas such as MT (Bradley et al., 1995), and MST (Maunsell & van Essen, 1983; Roy & Wurtz, 1990) are jointly sensitive to these cues.

However, deducing environmental structure is not the only task to which these signals could be put. When cues such as disparity and relative motion (RM) are available, their simultaneous presence may be diagnostic of the type of self-motion that gave rise to the dynamic visual information. Roy & Wurtz (1990) and Roy, Komatsu & Wurtz (1992) pointed out that under certain conditions the direction of self-motion could, in principle, be deduced from the particular combination of disparity and relative motion signals produced by movement through naturally occurring scenes. Consider figure 7.1. An observer travelling in the direction indicated by arrow A fixates in the middle distance of a slanting surface, in this case the ground plane. Retinally, objects lower in the visual field than the point of fixation translate in the direction opposite to the direction of motion (arrow B), whereas objects above fixation translate with the direction of travel (arrow C). Unfortunately, there is insufficient information in the motion signals to indicate the direction of self-motion if the depth structure is unknown.
Figure 7.1 As an observer travels in the direction indicated by arrow A, a typical landscape gives rise to opposing directions of motion on the retina, arrows B and C. When combined with depth information this relative motion (RM) is diagnostic of the observer's direction of travel. A scene containing a smooth change in depth produces a smooth speed gradient, with velocity falling to zero at the fixation point.
An observer travelling in the opposite direction looking at a slanted surface from below, such as a ceiling, would register exactly the same pattern of retinal motions. However, the addition of depth information disambiguates the meaning of the velocity gradient. Roy & Wurtz (1990) and Roy, Komatsu & Wurtz (1992) presented evidence for cells in primate area MST that have receptive field properties that take both depth and motion signals into account in their responses. These cells were found to have directional selectivity that depended on the sign of the disparity of the stimulus. For example, one cell would respond to rightwards motion only when disparity was crossed, but to leftwards motion only when disparity was uncrossed. The scenario illustrated in figure 7.1 produces a progressive velocity gradient (for a ground plane viewed at an oblique angle), with a singularity at the point of fixation. About this singularity the sign of retinal motion is reversed, producing a horizontally oriented shearing pattern (see figure 7.2). Retinal disparities follow the same pattern (not shown), with the largest disparities present at points furthest from fixation (in the vertical dimension), proceeding progressively to zero at fixation. Again, disparities on either side of fixation have opposite sign.

The fact that relative motion signals are of fundamental importance in gauging self-motion has been revealed in many studies of ‘vection’ (Lee & Lishman, 1975; Dijkstra, Schoner & Gielen, 1994; Palmisano, 1996; Palmisano, 2002). When visual signals induce an illusory perception of self-movement this is termed vection (Fischer & Kornmüller, 1930). As well as physiological evidence for a relationship between the coding of optic flow and vestibular signals (Duffy, 1998), there is psychophysical evidence in support of mechanisms that could code vection through the simultaneous availability of both depth and motion information (e.g. Nakamura & Shimojo, 1999; Nakamura & Shimojo, 2000). Unfortunately use of the vection phenomena to investigate RM and disparity in combination is not ideal as it is difficult to avoid using self-report as the dependent variable. However, several studies have addressed human sensitivity to the kind of relative motion signals shown in figure 7.2 using more objective measures, and some have combined disparity with similar forms of RM.
Figure 7.2 The retinal velocity gradient arising from the ego-motion depicted in figure 7.1. Arrows indicate the direction of motion, with the sign of the motion in the lower hemi-field being the inverse of that in the upper hemi-field. The length of each arrow is proportional to the speed component of the motion vector. The progressive change in arrow length with vertical displacement from the centre of fixation (black dot) reflects the smooth speed gradient observed when viewing a flat ground plane.

Motion detection thresholds have been shown to be lower where relative motion is present than when absent. Snowden (1992) compared sensitivity to unidirectional motion with sensitivity to relative motion in two-frame Random Dot Kinematograms (RDK's). The relative motion stimulus comprised two opposing hemi-fields of translating dots, forming a shear pattern similar to that shown in figure 7.2 (though lacking a speed gradient). Displacement thresholds
were established, showing that the human visual system is up to twice as sensitive to relative motion than to simple translation. A study by Golomb et al., (1985) investigated the sensitivity of both human and macaque observers to the spatio-temporal tuning of shearing motions, finding that the reaction time to detect shear displayed a U-shaped dependency on spatial frequency, and that both monkey and man displayed similar patterns of performance. Ono & Sato (2002) used an adaptation paradigm to investigate whether processing of shear stimuli could be attributed to specialised relative motion mechanisms. Observers were asked to discriminate the direction of a shearing stimulus after prolonged exposure to either uniform motion or shearing motion. Discrimination thresholds were elevated further when the adapting pattern was shear, leading the authors to suggest that a separate pathway exists for the processing of shearing motions. Other studies have also employed adaptation to show that the motion after-effect (MAE: an illusory motion perceived in a static image after prior exposure to a drifting pattern) is enhanced when relative motion is introduced in regions adjacent to the target region of motion (Ashida & Susami, 1997; Murakami & Shimojo, 1995).

Studies combining disparity information and relative motion signals come from a variety of investigators. Measuring the strength of vection in a self-report paradigm, Nakamura & Shimojo (1999) measured vection latency and duration experienced with two populations of randomly positioned moving dots. Disparity defined the dots at different depth planes, and the relative motions of the populations were manipulated. When one population was static, or moved in the opposite direction to the other, the dependent variables indicated that increased vection was perceived over alternative combinations of RM. Using a transparent motion Random Dot Kinematogram (RDK) Hibbard & Bradshaw (1999) found that the addition of binocular disparity to a global pattern of motion facilitated its detection when the opposing motions were assigned to different depth planes. Another study by Hibbard, Bradshaw & De Bruyn (1999) showed that discrimination of the direction of motion was aided when noise dots and signal dots were assigned different amounts of disparity.
In one of several related experiments Ziegler and Roy (1997) presented relative motions in distinct bands of dots travelling horizontally in opposing directions. Disparity was applied to all dots equally, defining them as either nearer or further in depth than fixation, the task being to discriminate which depth structure was being simulated. This arrangement was found to produce higher sensitivity for depth discrimination than a similar pattern containing unidirectional motion only, indicating an advantage for RM and disparity combined.

The purpose of the experiments described here was to ascertain whether the human visual system is uniquely responsive to combinations of relative motion and disparity depth cues that could be useful in calculating the direction of self-motion. One method for doing so is described by Graham (1989). A technique for revealing mechanisms that are specialised along more than one dimension of interest (‘analysers’) involves establishing the threshold level of stimulation required for a criterion level of performance when one cue is available, and comparing that with the threshold obtained when both cues are available. If the availability of two cues lowers the threshold then ‘summation’ is said to have occurred. To establish whether disparity information and relative motion combine to enhance sensitivity, motion coherence thresholds for detecting RM were measured in the presence of three forms of binocular disparity, indicating different depth structure. If thresholds were to be lower when the depth structure is consistent with a slanted surface then the case for the type of sensitivity recorded by Roy & Wurtz (1990) and Roy, Komatsu & Wurtz (1992) would be strengthened.
7.3 Experiment 1

7.3.1 Methods

Equipment

Stimuli were generated on an Apple Macintosh G4, 667MHz personal computer using software routines from the Video Toolbox, (Pelli, 1997). Images were rendered on a Philips FIMI Brightview high brightness monochrome display via an NVDA GeForce2 graphics card. The screen refresh rate was 67Hz, and the resolution was set to 640*480 pixels. With the screen measuring 37cm wide by 28cm deep, this gave approximately 17 pixels per cm.

A pair of stereo glasses (Stereographics, “CrystalEyes” EPC 2) was used to introduce binocular disparity. The glasses contain two lenses, one for each eye, which can be oscillated rapidly between transparent and opaque. Each lens was synchronised to alternate frames of the display raster, each of which showed one image from a stereo-pair. This arrangement resulted in a temporal frequency of 33Hz at each eye. Between trials the luminance was reduced as both lenses were transparent during this period.

Luminances:

At 67Hz the mean luminance through one lens of the stereo glasses was measured to be 44 cd/m².

Observers

Three volunteers participated (AS, SD and RW), including the author (RW). Each was an experienced psychophysics observer having normal visual acuity with corrective lenses worn where necessary. Stereopsis was assessed by asking subjects to discriminate between two orthogonally slanted surfaces defined solely by disparity. All observers performed in excess of 85% correct, where 50% represents chance performance.
**Stimuli**

A schematic representation of the stimuli is shown at figure 7.3. Two animated Random Dot Kinematograms (RDK’s) were presented, each in a randomly allocated square spatial window, surrounded by a solid black border to promote fusion. This border appeared prior to the first trial, and was present at all times during the block, as was the fixation cross, which was located halfway between the spatial windows. The black borders subtended 11.5*11.5 degrees of visual angle, and the fixation-cross subtended 1.9*1.9 degrees at the viewing distance of 57cm.

One RDK, the ‘signal’ animation containing shearing motion, comprised two sub-populations of dots. In one set each dot had its directional component assigned at random, and is consequently referred to as a noise dot. The other set had the sign of their directional components assigned according to their vertical position relative to centre of the window (which was in line with the fixation cross) in order to produce the shearing pattern depicted in figure 7.2. The speed of both sets of dots was linearly related to their vertical displacement in order to produce the speed gradient of figure 7.2. The direction of shear, ‘clockwise’ or ‘counter-clockwise’, was chosen at random on each trial. The alternative RDK, the noise animation, contained only dots with a randomly allocated directional component, though their distribution of speeds was also dictated by their vertical displacement. In this way the noise and signal animations differed in their directional distribution, but not their speed distribution.

There were three conditions of speed: nominally 5.8, 11.7 and 23.3 degrees/s. As the speed of each dot was scaled according to its vertical displacement from the centre of the dot field to produce a smooth shearing motion the median speed of the dots was 2.9, 5.8 and 11.7 degrees/s. Therefore dots moving at the nominal rate were only seen at the most extreme vertical edge of the window. The proportion of signal dots in the signal location was varied according to a Quest algorithm (Watson & Pelli, 1983).
Figure 7.3 Schematic representation of the three conditions of depth structure simulated using binocular disparity. Panel A shows the RDK elements receding smoothly in depth, as a landscape would ('slanted'). Panel B shows dots randomly distributed in depth ('random') and Panel C shows all dots at a constant depth ('fixed'). Distances shown were calculated with respect to the screen of the VDU, which was viewed from a
distance of 57cm. Filled dots are signal, whose directions produce shear motion (not shown), unfilled dots are directionally random noise dots.

**Disparity Conditions**

There were three levels of the disparity variable, with both animations being treated identically on a given trial. One level was a uniform disparity of 1.16 degrees applied to each dot, consistent with the dots occupying a flat plane perpendicular to the direction of the observers gaze ('fixed'). The second level had each dot assigned a random disparity between 1.16 degrees of crossed disparity and 1.16 degrees uncrossed, consistent with the dots being scattered in depth ('random'). The final level had a disparity gradient that varied with the vertical displacement of the dot ('slanted'). Below the centre of the window dots had crossed disparity, with a maximum of 1.16 degrees at the bottom of the window, and zero degrees at the centre. Above fixation disparities were uncrossed, with the maximum of 1.16 degrees at the top of the window. This arrangement was designed to be consistent with the depth profile of a flat plane viewed from above at an angle. Uncrossed disparity of 1.16 degrees simulated a dot placed 35cm beyond the screen, i.e. 92cm in total. Maximum crossed disparity simulated a dot approximately 16cm closer than the screen, or 41cm from the observer. Figure 7.1 shows schematic representations of the three conditions.

Each animated sequence was displayed for 1049mS. The dots used in the stimuli were difference of Gaussian elements (DoG’s), the spatial frequency content of which centred on 1.14 cycles per degree of visual angle at the viewing distance of 57cm employed. Fifty dots occupied each window, therefore the dot density was 2.83 dots per degree².

The Michelson contrast of the dots was 50% and the mean luminance of the display was 46 cd/m².

Each dot had a limited lifetime of 3 frames.
Procedure

Observers performed a block of practice trials to become familiar with the task, and to adapt to the low levels of luminance present. Each trial was separated by approximately 1 second as the computer calculated the animation for the next trial. All three levels of the disparity variable were interleaved pseudo-randomly within a block, with the constraint that an equal number of trials (50) were presented at each level. Each block of 150 trials lasted approximately 7 minutes, after which the subject could take a break if required. Each observer performed at least 4 blocks of trials for each level of the speed variable, resulting in a minimum of 200 data points for each unique combination of conditions. Speed conditions were explored in separate blocks of trials, their order being chosen at random.

7.3.2 Results

Figure 7.4 shows for each observer how coherence thresholds to detect shear motion, on the vertical axis, varied with the median speed of dots, on the horizontal axis. Three curves are shown, one for each condition of disparity. An effect of speed is evident, with thresholds diminishing with increasing speed. Statistically this was confirmed using a repeated measures ANOVA ($F_{(2,4)}=59.03, p<0.01$). Post hoc tests showed the significant differences to be between the 2.9 degrees/s condition ($p<0.01$) and 5.8 degrees/s; and between 2.9 degrees/s. and 11.7 degrees/s. ($p<0.01$). This is in accord with the Ziegler & Roy (1997) who found that sensitivity to depth structure in the presence of relative motion was enhanced by speed. ANOVA also confirmed that while there was no interaction between speed and depth, there was an effect of depth structure (repeated measures, $F_{(2,4)}=8.4, p<0.05$), which post hoc testing showed was between slanted and random conditions ($p<0.05$), with slanted condition obtaining the higher thresholds. This is contrary to what might be expected if speed gradient and disparity information were to combine in an extra-ordinary way when both were generated by motion through a smoothly receding landscape.
Figure 7.4 Motion coherence thresholds for the detection of shear motion (vertical axis) as a function of median speed (horizontal axis) and disparity condition (red, blue and yellow curves). The first three graphs show data for each of the three observers separately. Error bars are 95% confidence intervals. As the data for each subject were similar the final graph shows the data for all three observers collapsed. Error bars are one standard error of the mean.

However, some aspects of the stimulus used in Experiment 1 may not have been conducive to stimulation of mechanisms that are optimised for optic flow information obtained from realistic scenes. One concern centred on the possible impoverishment of the disparity signal, given the relatively short lifetime of the dots (3 frames). The experiment was repeated with an increased dot lifetime of 6 frames (see figure 7.5). The increase in lifetime did change the results slightly, with no main effect of disparity type found (repeated measures ANOVA, $F_{(2,4)} = 3.7$, $p>0.05$).

This effect of lifetime is in disagreement with the findings of Treue, Husain & Andersen (1991) who showed that the build up of structure from motion using RM and disparity to define a rotating cylinder was not improved with dot lifetimes in excess of circa 125mS. In the previous experiment a lifetime of 3 corresponded to a duration of 180mS, and in the present experiment the lifetime of 6 equated to 360mS, both above the level Treue, Husain & Andersen (1991) found to correspond to asymptotic performance.
Once again a main effect of speed is present ($F_{(2,4)} = 7.7, p<0.05$), confirming the general pattern of results from Experiment 1, with no interaction between speed and disparity being recorded.

![Figure 7.5 Lifetime=6 frames. Motion coherence thresholds for the detection of shear motion (vertical axis) as a function of median speed (horizontal axis) and disparity condition (red, blue and yellow curves). As the data for each subject were similar the graph shows the data for all three observers collapsed. Error bars are one standard error of the mean.]

**7.4 Experiment 2**

In a study investigating how relative motion and disparity combine, Ziegler & Roy (1998) found that high retinal speeds and large values of disparity in the periphery were effective in enhancing observers’ ability to discriminate whether stimuli were nearer or further in depth than the fixation plane. They identified that optimum speed was around 60 degrees/s, in the presence of over 4 degrees of retinal disparity. As the values of speed and disparity parameters in Experiment 1 were substantially below those of Ziegler & Roy (1998) they were increased for Experiment 2. Another consideration in the perception of optic flow is the spatial extent over which the stimulus is presented, as many of MSTd receptive fields have shown that they cover wide areas of visual space, often in excess of 40 degrees (Tanaka & Saito, 1989; Duffy & Wurtz, 1991). As a
consequence of the way dot speeds were calculated in these experiments enlarging the animation window in which the RDK was shown had the desired effect of increasing dot speeds throughout the window, as well as providing a wider field of stimulation. However, this increase in window size also meant that only one window could be shown at a time (for computational reasons), so Experiment 2 employed a two temporal interval, forced choice procedure. Further dot speed enhancement was achieved by increasing the refresh rate.

7.4.1 Methods
As for Experiment 1, with the following exceptions:
The display refresh rate was increased to 120Hz, exposing each eye to a flicker rate of 60Hz. At this setting the luminance of the display was 51 cd/m². The higher refresh rate reduced the duration of the animation to 583mS. The animation window was increased in size to 16.9 * 16.9 degrees, with the fixation cross remaining at 1.9 * 1.9 degrees. The increased size of the windows precluded both animations being shown on screen together. Therefore they were shown in two temporal intervals, their order being randomised. Button ‘1’ on the alphanumeric keyboard was used to indicate that shear motion was seen in the first interval, button ‘3’ if seen in the second interval.

Once again there were three speed conditions in this experiment. However, speeds were increased by two factors. The elevated refresh rate caused the animation to be shown in a shorter period of time, therefore increasing dot speed proportionately. In addition the enlargement of the animation window increased mean velocity, as each dot’s speed was a function of its vertical displacement from the window’s centre. Peak speeds at the vertical extremities of the windows were: 15.7, 31.4 and 62.5 degrees/s, spanning the range investigated by Ziegler & Roy (1998). The median speeds were therefore: 7.8, 15.7 and 31.4 degrees/s. The disparity was increased from the nominal 1.16 degrees to 2.6 degrees. This level is less than the 4 degrees used by Ziegler & Roy (1998), but it was found that the perception of depth broke down at greater disparities.
Three observers took part, though one change to the personnel of the previous experiments was necessary, MC taking the place of SD.

7.4.2 Results

Figure 7.6 shows the threshold performance for each of the three observers, separately. The general pattern for AS and RW is similar to that seen in previous experiments, with little difference between disparity conditions. MC's data appear anomalous in comparison to these results, and those recorded previously, showing an interaction between speed and disparity at the highest (11.7 degrees/s) median speed condition. In the case of the two more experienced observers there is no evidence in the data that the manipulations to display size, dot speed and disparity have significantly altered the ability to process relative motion and disparity together.

One possible reason for the absence of an effect may lie in the way that noise dots in the interval containing relative motion were distributed in depth. As figure 7.3 shows, both signal and noise dots obeyed the disparity parameter of the depth condition being explored, i.e. if a slanted depth plane was called for then both shearing dots and those with randomised directions inhabited the virtual slant. Thus noise dots were 'intermingled' with signal dots in depth, and may have served as a dual mask within mechanisms sensitive to both parameters. Two recent studies suggest that segmentation of motion signals can be enhanced if they are differentiated not just by direction, but also by disparity.
Figure 7.6 Motion coherence thresholds for the detection of shear motion (vertical axis) as a function of median speed (horizontal axis) and disparity condition (red, blue and yellow curves). The three graphs show data for each of the observers separately. Error bars are 95% confidence intervals.
Hibbard, Bradshaw & De Bruyn (1999) presented subjects with a similar motion coherence task to that used here, except observers were asked to discriminate between opposite directions of motion. In a series of three experiments these authors showed that separating noise and signal dots into distinct depth regions lowered coherence thresholds, compared to intermingling signal and noise dots throughout a volume. Snowden & Rossiter (1999) also found that separating noise and signal dots in depth was effective in lowering motion coherence thresholds in a direction discrimination task. Experiment 3 was conducted to establish whether the mixing of signal and noise disparities had prevented any advantage for the combination of relative motion and congruent depth information from being revealed.

7.5 Experiment 3

7.5.1 Methods

Separating the noise and signal dots into distinct depths meant that a detection task was no longer suitable as disparity provided a redundant cue for discriminating the signal-plus-noise interval from the noise-alone interval. Therefore the task set for Experiment 3 was direction discrimination, in which disparity is uninformative, though it may still be effective in influencing the motion coherence threshold (Hibbard, Bradshaw & De Bruyn, 1999; Snowden & Rossiter, 1999). Consequently one spatial location was presented with the direction of shear (immersed in noise) being chosen at random. Subjects indicated counter-clockwise shear (upper half of the animation translating left, lower half translating right) by pressing button ‘1’, and clockwise shear using button ‘3’. Motion coherence thresholds were established, again using Quest (Watson & Pelli, 1983) to vary the proportion of signal dots in the animation. Signal dots were assigned depth structure according to the three disparity conditions used previously (see figure 7.7). Noise dots, however, were assigned a fixed disparity of 1.16 degrees (uncrossed) throughout each trial, regardless of the distribution of the signal dots. This had the effect of placing all the noise dots
at a distance of 35 cm behind the point of fixation. The noise dots were placed at
greater depth than the signal dots as Hibbard, Bradshaw & De Bruyn (1999)
showed that an asymmetry may exist in the processing of signal and noise dots
as a function of depth ordering. When noise dots were placed in front of signal
dots a distinct elevation in motion coherence thresholds was obtained for two of
their four observers.

For efficiency it was decided to restrict the speed of relative motion to the
medium-speed condition (median speed of 5.8 degrees/s).

Observers were AS, SD and the author, RW, as in Experiment 1.

Other methodological details were the same as for the previous experiment.
Figure 7.7 Schematic representation of the three conditions of depth structure simulated in Experiment 3. Signal dots (filled) obey the three disparity conditions, as before. Noise dots now occupy their own depth plane at all times.
7.6 Results and Discussion

Figure 7.8 shows how motion coherence thresholds to discriminate the direction of shear motion varied with the depth structure variable for each of the three observers separately. Although absolute thresholds differ between subjects, with SD requiring substantially lower signal than the others, the results from all three observers are consistent: no condition of disparity was effective in influencing the threshold for direction discrimination. Although the ‘fixed’ condition was intended as a control condition in these experiments, its use in Experiment 3 mimicked the condition in the experiments of Snowden & Rossiter (1999) and Hibbard, Bradshaw & De Bruyn (1999) that produced a strong effect of disparity, i.e. two distinct depth planes (see figure 7.7). These results were surprising in the light of direction discrimination results obtained in two independent studies (Hibbard, Bradshaw & De Bruyn, 1999; Snowden & Rossiter, 1999), which showed a clear effect of separating signal and noise dots using binocular disparity. Although neither study contained a condition like the ‘slanted’ depth structure used here, the effects those authors reported seem to be contingent on a separation of signal and noise dots into two distinct planes, which would explain a lack of an effect in the ‘slanted’ condition. The study by Snowden & Rossiter (1999) compared thresholds when the dots were arranged in separate depth planes with the case when they were all assigned zero disparity. Of the conditions explored by Hibbard, Bradshaw & De Bruyn (1999) separation of signal and noise dots into two planes was effective in lowering thresholds, compared to an arrangement where dots were distributed in depth. These results are difficult to reconcile, even by appealing to the main difference between those studies and this: the presence of spatially separated relative motion, i.e. shear. In a separate study Hibbard & Bradshaw (1999) showed that the detection of transparent motion benefited from the application of disparity to segregate the opposing populations into two depth planes, so relative motion per se does not seem to be a good candidate to explain the lack of an effect.
Figure 7.8 Motion coherence thresholds for the discrimination of shear direction. Thresholds are shown on the y-axis, as a function of disparity type on the x-axis, for each of the three observers separately.

Other than the use of shear there were differences between the experiments. For example, two-frame Random Dot Kinematogram’s were used by the previous authors, stimuli were somewhat smaller in their experiments, and dot density was generally higher. None of these parameters seems likely to account for the differences, without concluding that the effects found were highly particular to the stimulus conditions used. Perhaps the most significant departure between those experiments and this one was in the amount of binocular disparity applied to the dots. The disparities used in the present experiments were significantly greater (Experiments 1 & 2: 1.16 degrees crossed to 1.16 degrees uncrossed; Experiment 3: 2.6 degrees crossed to 2.6 degrees uncrossed), in keeping with those used by Roy, Komatsu & Wurtz (1992) (3 degrees crossed to 3 degrees uncrossed). Hibbard & Bradshaw (1999) separated their opposing transparent populations by 16 min of arc, corresponding to 0.267 degrees of visual angle. Snowden & Rossiter (1999) used quite small disparities of 0.048 degrees crossed
and uncrossed to separate their dot populations. In their experiments Hibbard, Bradshaw & De Bruyn (1999) did not use more than 40 minutes of arc (0.667 degrees), and found that 10 min of arc separation between noise and signal was optimal for discrimination. In their discussion, the latter authors attribute improved discrimination performance to an attentional strategy, observers parsing the depth structure in order to segment the populations prior to direction judgements. However, none of the conditions presented here incorporated a zero disparity condition, so a simple disparity segmentation strategy was not tenable. Indeed the motivation for this work was not an interest in segmentation cues, but in summation. Evidently no comparative improvement in direction discrimination is possible when all conditions contain depth structure, suggesting that any mechanisms sensitive to RM and disparity are either not sufficiently prevalent to influence detection and discrimination performance, or are not so tightly tuned as to prefer one depth structure over another.

Binocular disparity is a feature of visual processing that has traditionally been regarded as being most relevant over short distances. The image presented to each eye becomes essentially the same when viewing distant objects because of the comparatively small lateral displacement of the eyes in the head, i.e. there is no disparity to exploit. However the relative motion signals arising from movement through depth remain effective at large viewing distances, leading to the question: is it advantageous for biological systems to combine disparity and motion parallax? From a theoretical point of view the answer is probably yes, since disparity and motion parallax are both effective at short distances. Indeed a 6cm movement of a monocular observer generates the same amount of disparity as is seen by a stationary observer having the benefit of binocular vision (Helmholtz, 1909; Rogers & Graham, 1982). Practically too, experiments in both humans (Hibbard, Bradshaw & Hibbard, 1999; Snowden, & Rossiter, 1999) and monkey (Bradley et al., 1995; Cao & Schiller, 2002; Roy, Komatsu & Wurtz, 1992; Roy & Wurtz, 1990) suggest that there are interactions between these cues under certain conditions. However the fact that no advantage was
found here when both cues were made available here may reflect the circumstances under which there is no value in combining them.
Chapter 8
Conclusions

The experiments reported in this thesis sought to explore the sensitivity of complex motion mechanisms along some of the many dimensions to which they respond. Established techniques were used to investigate questions that have received only a little attention in recent years, and new techniques were applied to re-examine issues that have been addressed by previous authors, correcting for methodological shortcomings. For example, in chapter 3 the range of directions to which translation, rotation and radial mechanisms respond was examined using a conventional technique that has previously been applied only to unidirectional motion. The results suggested that directional bandwidths are narrower than previously thought. Chapter 4 concerned the spatial frequency selectiveness of mechanisms that respond to radial motion, a retinal pattern that is produced when a moving observer fixates on his direction of heading. Spatial filtering of random dot elements allowed the dissociation of size and spatial frequency. This led to conclusions concerning the preferred spectral content of motion detectors, which was found to vary with speed. Chapter 5 explored whether radial patterns were indeed interpreted as motion in depth via the introduction of conflicting depth cues in the form of binocular disparity. As the frequently reported perceived speed enhancement that radial patterns enjoy was not affected by this manipulation it was concluded that there was no evidence to suggest that observers' visual systems interpreted radially expanding patterns as motion in depth. The speed tuning of translation, rotation and radial mechanisms was examined using a new technique in chapter 6. Careful balancing of the speeds of complex motion dots and directionally random noise dots lead to quite high speed discrimination thresholds that, while receiving some support from the existing literature, were deemed to be unreliable. Consequently a different and previously used technique was applied for the first time to the measurement of speed discrimination in rotating patterns. Discrimination thresholds more in keeping with existing data on discrimination in translating patterns were found,
and thresholds diminished with increasing speed reflecting a failure of Webers law. Finally, chapter 7 investigated whether a special sensitivity existed for combinations of binocular disparity and relative motion that are diagnostic of an observer’s direction of travel. In contrast to related work no performance advantage was obtained when these cues were both present, suggesting that joint sensitivity to relative motion and binocular disparity is not a prominent source of self-motion perception.

Future directions
Direction bandwidths. The measurement of direction bandwidths in chapter 3 relied heavily on two assumptions. Firstly, the width of mechanisms measured using the signal in noise techniques (#1: low contrast dots; #2: high contrast dots immersed in noise dots) was taken to be equal to 6 standard deviations of a Gaussian shaped response profile. As was pointed out in that chapter the asymptotic nature of the Gaussian profile makes measurement of a mechanism at its base a problematic task. It is reasonable to assume that such a mechanism does not continue to contribute to the response of the system at infinitesimally small response levels, but just where it ceases to participate is unknown. Secondly, the degree of interaction between neighbouring mechanisms was assumed to be zero in the case of low contrast and high contrast dot stimuli. This assumption follows the logic of the labelled lines hypothesis (at threshold only the most sensitive mechanism responds), and was necessary as a deduction concerning the half height width of a mechanism was only possible if only one mechanism was responsible for the observers’ responses. Both these assumptions were required because of the attempt to measure mechanism width at its base. To avoid these difficulties future attempts to measure mechanism widths would do better to seek the maximum responsiveness of a mechanism and describe its loss of activity as the stimulus is varied to sub-optimal values. The half-height points of the mechanism are then explicitly available, as is its response profile, and the assumptions made in the current study are obviated.
In the experiments of chapter 3 only bandwidths of mechanisms centred on the cardinal directions of optic flow (rotation and radial motion) were examined. The finding that these are narrow, supporting the assertion that mechanisms tuned to intermediate patterns may exist, suggests that attempts to measure intermediate directions of flow is a necessary next step. Even if rotation and radial mechanisms had been found to have similar widths, it would not follow that intermediate mechanisms would have similar direction bandwidths, and a fuller understanding of the operation of the motion system would result from measuring their properties.

The spatial frequency preferences of radial motion detectors in chapter 4 provides additional support for the idea that local motion detectors form the basis of complex motion processing, donating spatial characteristics that constrain the range of responsiveness of complex mechanisms. One proposal arising from those experiments, with equivocal support from another laboratory, is that simulations of high-speed motion in depth are optimally detected if spatial scale is increased between samples of a dots' trajectory. As radial expansion was the only pattern examined in those experiments it is not known whether similar results would be obtained with radially contracting stimuli. Radial contraction is a pattern that is much less prevalent in the normal experience of visually guided organisms than radial expansion as it requires the observer to retreat without looking where they are going. This issue could easily be explored as an extension to the present work. As for rotational patterns, it is not known whether any spatial frequency specificity would be evident, and any such finding would be of interest given the lack of motion in depth (and hence size and spatial frequency change) implied by a rotating pattern. Intermediate patterns, such as spirals, do have a valid interpretation as a pattern arising from motion in depth, and represent another area in which spatial scale specificity is of interest. The detection paradigm used in chapter 4, while capable of illuminating the sensitivity of visual mechanisms by teasing out small differences in performance, does not necessarily indicate which parts of the spatial spectrum
are used in visually guided behaviours. Further studies could explore how spatially impoverished environments influence ecologically motivated behaviours, such as navigation.

In chapter 5 it was found that the exaggerated perceived speed of radially expanding patterns was not influenced by motion in depth cues such as binocular disparity and size change. This does not constitute evidence against radial motion being interpreted as motion in depth and for future work many other depth cues are available to influence the interpretation of such patterns. However, any attempts to do so may fail simply due to the salience of the relative motion cue. Attempts to null the speed illusion are, therefore, one-sided with no ability to draw a strong conclusion from a negative result. Clearly depth cues could be added that concur with the motion signalled by radial expansion, rather than in conflict with it. However such manipulations would not be expected to further amplify the illusion, but merely render the pattern less ambiguous (i.e. indicating that it is not due to a deformation at constant depth). Nevertheless, attributing the speed illusion to a motion in depth interpretation would make available a useful psychophysical tool. The speed illusion is a phenomenon that could then be used as a diagnostic test of the sensation of motion in the third dimension, a perception that is entirely subjective, and that could benefit from the availability of a measure that is tangentially related to the perception of motion in depth.

As the results of the novel technique used in chapter 6 speed were rejected discrimination thresholds were only measured for rotation patterns with the alternative methodology. It remains to be determined what the speed increment sensitivity of radial motion detectors is when local motion signals are removed or rendered unreliable cues. The ecological significance of navigation in a hazardous environment suggests that speed tuning in radial motion detectors ought to be refined. Unfortunately a suitable technique has not yet been devised to answer this question. Attempts to present radial motion in annuli of differing
sizes (as with the rotating patterns) would lead to ambiguity concerning whether both stimuli inhabited the same depth plane, and hence what their metric speed could be. Making use of the speed illusion investigated in chapter 5 is ruled out on two counts: it has not yet been shown that it can be used as diagnostic of motion in depth, and it has been found to be rather too noisy a dependent variable to use for accurate measurements. Until a technique can be devised to emasculate local speed cues effectively this question must remain an open one.

In spite of evidence from previous studies to the contrary, the final chapter showed that stereo cues did not augment performance in the detection or discrimination of opposing relative motion (shear). As such, no evidence for mechanisms exhibiting a strong degree of dual tuning was found, despite the theoretical utility of such a combination in the perception of self-movement. This is surprising given the inevitable connection between retinal motion signals and the depth structure of the environment that determines their relative magnitudes. Although it was not specifically the aim of the experiments to generate or measure visually induced motion (vection) future work exploring combinations of relative motion and other depth cues such as binocular disparity might benefit from that literature. Vection is most prominent when very large visual regions are stimulated, and when extended stimulus duration allow the perception of self-movement to build up. Under these conditions enhanced sensitivity to combinations of motion and depth cues may be revealed, while retaining non-subjective forced choice psychophysical methods.
Bibliography


Harris, M.G., & Giachritsis, C.D., 2000. "Course-grained information dominates fine-grained information in judgments of time-to-contact from retinal flow". *Vision Research*, 40, 601-611.


Hibbard, P.B., Bradshaw, M.F., & De Bruyn, B., 1999 “Global motion processing is not tuned for binocular disparity”. Vision Research, 39, 961-974.


Newsome, W.T., & Pare, E.B., 1988."A selective impairment of motion perception following lesions of the middle temporal visual area (MT)”. Journal of Neuroscience, 8, 2201-2211.


### Appendix 1

List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANOVA</td>
<td>Analysis of Variance</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence Interval</td>
</tr>
<tr>
<td>CPU</td>
<td>Central Processing Unit</td>
</tr>
<tr>
<td>CRT</td>
<td>Cathode Ray Tube</td>
</tr>
<tr>
<td>DOG</td>
<td>Difference of Gaussian</td>
</tr>
<tr>
<td>DRDS</td>
<td>Dynamic Random Dot Stereogram</td>
</tr>
<tr>
<td>ERD</td>
<td>Elaborated Reichardt Detector</td>
</tr>
<tr>
<td>FOE</td>
<td>Focus Of Expansion</td>
</tr>
<tr>
<td>KDE</td>
<td>Kinetic Depth Effect</td>
</tr>
<tr>
<td>LGN</td>
<td>Lateral Geniculate Nucleus</td>
</tr>
<tr>
<td>MAE</td>
<td>Motion After Effect</td>
</tr>
<tr>
<td>MST</td>
<td>Medial Superior Temporal (area)</td>
</tr>
<tr>
<td>MT</td>
<td>Middle Temporal (area)</td>
</tr>
<tr>
<td>OS</td>
<td>Operating System</td>
</tr>
<tr>
<td>PDF</td>
<td>Probability Density Function</td>
</tr>
<tr>
<td>PSE</td>
<td>Point of Subjective Equality</td>
</tr>
<tr>
<td>RDK</td>
<td>Random Dot Kinematogram</td>
</tr>
<tr>
<td>RF</td>
<td>Receptive Field</td>
</tr>
<tr>
<td>RGB</td>
<td>Red Green Blue</td>
</tr>
<tr>
<td>RM</td>
<td>Relative Motion</td>
</tr>
<tr>
<td>TTC</td>
<td>Time To Contact</td>
</tr>
<tr>
<td>VDU</td>
<td>Visual Display Unit</td>
</tr>
<tr>
<td>WRT</td>
<td>With Respect To</td>
</tr>
<tr>
<td>2AFC</td>
<td>Two Alternative Forced Choice</td>
</tr>
</tbody>
</table>