Adaptive Motion Analysis in Machine and Biological Vision

Colin W.G. Clifford
University College London

December 1996

THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
Abstract

This thesis investigates the problem of machine and biological motion perception from an adaptive systems perspective. Adaptive temporal filters are incorporated into two established image motion analysis algorithms, one each from machine and biological vision. The correlation-based Reichardt detector (Reichardt, 1961) is equipped with adaptive filters to account for existing electrophysiological data on motion adaptation in the insect lobula plate. The wider applicability of the model is tested by recording electrophysiologically from cells in the mammalian nucleus of the optic tract (NOT), and by investigating rapid adaptation to motion in human psychophysical observers. Adaptive temporal filters are also used to implement a phase-based scheme for image velocity measurement (Fleet & Jepson, 1989). The use of adaptive filters reduces the computational load of the phase-based scheme while maintaining performance on a synthetic test sequence. The adaptive scheme shows an advantage over its non-adaptive counterpart at high levels of noise as adaptation serves to maximise the signal power in the output of the filters.
Acknowledgements

I am grateful to many people at UCL, especially Chris Benton, Oliver Braddick, Kate Bradford, Willie Curran, John Draper, Paul Hibbard, Alan Johnston, Peter McOwan, Justin O'Brien and Simon Prince for their advice, assistance, tolerance and general calming influence over all or part of the last three years, my long-suffering housemates, Paul Chambers and Jim Davy, for their company and friendship, the Psychology Allstars football team for many hours of cold, muddy enjoyment, and Keith Langley for his enthusiastic supervision.

I would also like to thank Johannes Zanker for inviting me to visit ANU, Michael Ibbotson and Richard Mark for introducing me to electrophysiology, and everyone who helped make my three months in Canberra so enjoyable, especially the Zanker family and Cathy Leamey.

Finally, I would like to thank my mother, Moira Lee, for the continued faith she has shown in me and support she has given me over the years.
This thesis is dedicated to the memory of my father,

Publications arising from the thesis:

Chapter 2:


Chapter 3:


Chapter 4:


Chapter 5:

Contents

1 Introduction to image motion analysis ............................................................... 14
   Motion as a fundamental visual dimension ................................................. 15
   The use of motion information ............................................................... 17
   Existing theories of motion perception ..................................................... 18
      Intensity-based approaches ............................................................... 19
      Relationships between approaches .................................................. 22
   Methodological issues in the study of biological vision ......................... 23
      Computational vision as a Lakatosian research programme .............. 25
   Adaptive motion analysis in biological vision ........................................... 27
   Adaptive motion analysis in machine vision ............................................ 28
      Steering or shearing? ................................................................. 31
      Recursive filters ........................................................................ 33
   Overview ........................................................................- .................... 34

2 An adaptive Reichardt detector model of temporal adaptation in fly
   motion vision ......................................................................................... 36
   The Reichardt detector ........................................................................ 37
   Adaptation in fly motion analysis .......................................................... 40
      The phenomenology of adaptation .................................................. 40
      The origin of the adapting signal ................................................... 41
      Establishing the functional significance of adaptation ................. 42
3 Extending the model to mammalian motion adaptation ........................................... 78

Introduction ........................................................................................................... 79

Model ...................................................................................................................... 80
4 Psychophysics of motion adaptation parallels insect and mammalian electrophysiology

Introduction ....................................................... 107
Method ................................................................. 108
Results ................................................................. 111
Perceived speed during adaptation ..................... 111
Differential velocity sensitivity during adaptation .... 117
Discussion ............................................................ 120
Summary .............................................................. 121
5 An application of adaptive filters to the measurement of image velocity in 

<table>
<thead>
<tr>
<th>machine vision</th>
<th>122</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>123</td>
</tr>
<tr>
<td>Measuring velocity from band-pass filters</td>
<td>125</td>
</tr>
<tr>
<td>Calculating confidence estimates</td>
<td>130</td>
</tr>
<tr>
<td>Spatial filters</td>
<td>130</td>
</tr>
<tr>
<td>Space-time oriented filters</td>
<td>132</td>
</tr>
<tr>
<td>Causal recursive temporal filters</td>
<td>132</td>
</tr>
<tr>
<td>Cascaded temporal filters</td>
<td>135</td>
</tr>
<tr>
<td>Digital filter design</td>
<td>136</td>
</tr>
<tr>
<td>Temporal adaptation</td>
<td>139</td>
</tr>
<tr>
<td>Dynamics of adaptation</td>
<td>143</td>
</tr>
<tr>
<td>Results</td>
<td>144</td>
</tr>
<tr>
<td>Implementation</td>
<td>144</td>
</tr>
<tr>
<td>Error measures</td>
<td>145</td>
</tr>
<tr>
<td>Measured flow fields</td>
<td>146</td>
</tr>
<tr>
<td>Conclusion</td>
<td>150</td>
</tr>
</tbody>
</table>

6 Conclusions | 152 |
<p>| Rapid neural adaptation to motion | 154 |
| Motion adaptation in insects | 154 |
| Comparing motion adaptation in insects and mammals | 155 |
| Rapid adaptation to motion in humans | 156 |
| Adaptive temporal filters for the phase-based measurement of image | 157 |</p>
<table>
<thead>
<tr>
<th>Number</th>
<th>Figure Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Strategies for the detection of motion</td>
<td>20</td>
</tr>
<tr>
<td>1.2</td>
<td>Motion as orientation in space-time</td>
<td>21</td>
</tr>
<tr>
<td>1.3</td>
<td>The horizontal cells of the fly lobula plate</td>
<td>29</td>
</tr>
<tr>
<td>1.4</td>
<td>Neural pathways to the nucleus of the optic tract</td>
<td>30</td>
</tr>
<tr>
<td>1.5</td>
<td>Motion in the frequency domain</td>
<td>32</td>
</tr>
<tr>
<td>2.1</td>
<td>The Reichardt detector</td>
<td>39</td>
</tr>
<tr>
<td>2.2</td>
<td>The adaptive Reichardt detector</td>
<td>43</td>
</tr>
<tr>
<td>2.3</td>
<td>Amplitude spectra of Reichardt detector arrays</td>
<td>45</td>
</tr>
<tr>
<td>2.4</td>
<td>Response of HI and model to maintained motion</td>
<td>51</td>
</tr>
<tr>
<td>2.5</td>
<td>Time constant as a function of temporal frequency of adapting stimulus</td>
<td>54</td>
</tr>
<tr>
<td>2.6</td>
<td>Response recovery after adaptation</td>
<td>57</td>
</tr>
<tr>
<td>2.7</td>
<td>Response to increments and decrements of velocity around baseline</td>
<td>59</td>
</tr>
<tr>
<td>2.8</td>
<td>Response to velocity modulated stimuli</td>
<td>62</td>
</tr>
<tr>
<td>2.9</td>
<td>Response modulation as a function of frequency of velocity modulation</td>
<td>64</td>
</tr>
<tr>
<td>2.10</td>
<td>Response of HS and model to maintained motion</td>
<td>66</td>
</tr>
<tr>
<td>3.1</td>
<td>Old and new versions of adaptive Reichardt detector</td>
<td>82</td>
</tr>
<tr>
<td>3.2</td>
<td>Response to sustained motion at each of two temporal frequencies</td>
<td>85</td>
</tr>
<tr>
<td>3.3</td>
<td>Response to motion of wallaby NOT stimulated through narrow aperture</td>
<td>86</td>
</tr>
<tr>
<td>3.4</td>
<td>Time constant as a function of temporal frequency of adapting stimulus</td>
<td>88</td>
</tr>
</tbody>
</table>
3.5 Adapted impulse responses of wallaby NOT neuron .................................. 89
3.6 Time constants after adaptation in four conditions .............................. 91
3.7 Temporal frequency response functions before and after adaptation ... 93
3.8 The model's impulse response before and after adaptation ............... 95
3.9 The locus of adaptation in wallaby NOT cells and the model ............. 97
3.10 Recursive implementation of the temporal filter used in the model ... 104

4.1 Experimental procedure used for psychophysics ............................... 110
4.2 Perceived speed as a function of adaptation duration ....................... 113
4.3 Controlling for contrast adaptation .................................................. 115
4.4 Further controls for perceived speed experiment ............................. 116
4.5 Adaptation of perceived speed varying spatial and temporal frequencies 118
4.6 Detection of velocity modulation as a function of adaptation duration 119

5.1 Adaptive scheme for image velocity measurement ............................ 126
5.2 Spatial filters used to implement the phase-based scheme ............... 131
5.3 The construction of spatiotemporally oriented filters ....................... 133
5.4 Temporal filters used to implement the phase-based scheme ............ 137
5.5 Recursive implementation of the temporal filters ............................ 140
5.6 Filters adapting to a plane of power in frequency space .................. 142
5.7 Confidence map from the diverging tree sequence .......................... 147
5.8 Correct and measured motion fields from the diverging tree sequence 148
5.9 Measured flow field from the Rubik cube sequence ........................ 151
List of tables

5.1 Parameters for the discrete-time transfer function of the third-order filter and its derivative according to equation 5.19 ........................................... 139

5.2 Mean errors obtained from frame 37 of the diverging tree sequence, at various levels of added noise, using fixed and adaptive filters ............... 149
Chapter 1

Introduction to image motion analysis
Next to the simple detection of light and dark, the ability to perceive motion may
be the oldest and most basic of visual capabilities (Young & Lesperance, 1993).
The world that we live in is constantly in motion and, when awake, we are
constantly moving around interacting with objects and other organisms. Our view
of the world is constantly changing. This means that the ecological stimulus for
vision is inherently spatiotemporal: a changing optic array (Lee, 1980).
Consequently, we might expect motion detection to be a fundamental role of any
biological vision system. This does indeed seem to be the case: while numerous
animals lack colour vision or significant binocular vision, the ability to see motion
is common among all visual animals (Nakayama, 1985).

Visual analysis is an active process involving the continual adaptation of
elementary processing units. Such rapid adaptation is a fundamental property of
vision with moment to moment relevance to perception. While much previous
work has considered the functional benefits of photoreceptor adaptation (e.g.
Laughlin, 1989), adaptation at subsequent levels of the visual system has often
been neglected or viewed as a limitation of the system associated with neural
fatigue. In this thesis, functional roles for rapid adaptation in the robust and
efficient processing of motion information are considered in the contexts of
machine and biological vision.

**Motion as a fundamental visual dimension**

In studying motion processing separately from other visual functions we assume a
high degree of modularity within the visual system such that we can talk of a specific motion sub-system (Ullman, 1991). Below, the evidence for functional specialisation within the primate visual system is reviewed, and for motion as a fundamental visual dimension, like form and colour, rather than the result of a secondary stage of processing based on prior extraction of some other attribute.

The first demonstration of a separate motion process was provided by the "waterfall illusion" (Wohlgemuth, 1911). The waterfall illusion is an example of the motion aftereffect (MAE) whereby stationary objects are seen to move in the direction opposite to the motion of a previously fixated object, giving rise to a perceptual dissociation of motion and position. This is a paradox unless motion and position are regarded as separate sensory dimensions (Gregory, 1966). More recently, studies of the cytoarchitecture of the visual cortex coupled with single cell recording have provided evidence for functional specialisation at the neural level (Zeki & Shipp, 1988; Zeki, 1993). In the macaque monkey, the neural substrate of motion perception is a pathway originating in the M ganglion cells in the retina, which relay to the magnocellular layers of the lateral geniculate nucleus (LGN). From the LGN the output is relayed to layer 4B of the primary visual cortex, V1, and from there to V5, both directly and through the thick stripes of area V2 (Zeki, 1993). Corroboration of functional specialisation at the neural level in the human visual system has come from a clinical study by Zihl et al reporting a case of akinetopsia, a syndrome in which a patient is specifically unable to perceive motion following a lesion to the extra-striate cortex (Zihl et al, 1983), and from PET studies leading to the location of human V5 (Zeki et al, 1991).
The use of motion information

The previous section stressed the fundamental role of motion processing in visual function. This point is illustrated by examining various functions which rely on motion information. It is helpful to make a distinction between the low-level processes which extract motion information from the optic array and higher-level processes contingent upon that motion information. Transduction of light incident on the retina gives rise to a time-varying pattern of activity. The pattern of retinal activity is the result of the projection of a spatially three-dimensional environment onto the two-dimensional retina. The projection of the 3-D velocity field onto the retina is known as the optical flow field (Horn & Schunk, 1981). It is the role of low-level motion processing to provide an estimate of the optical flow field for the higher level functions reviewed below.

Given a 2-D retinal image, the visual system must provide the third dimension, depth. One of the most fundamental monocular cues to depth is motion parallax. The optical flow field contains a wealth of information regarding the layout of the surrounding environment (Longuet-Higgins & Prazdny, 1980), as is evidenced from our ability to achieve a perception of depth from moving random dot displays (Rogers & Graham, 1982) similar to that seen in random dot stereograms (Julesz, 1971). However, while relative depth information is available solely from motion cues, absolute distances and absolute velocities are not uniquely determined by the optical flow field since a fast, distant motion and a slow, near motion can produce identical projections onto the retina. What is uniquely
determined by the optical flow field is the time-to-contact, $\tau$, between the object and the observer, which from an ecological standpoint is a more basic parameter (Lee, 1980).

Depth and time-to-contact information are vital if we are to navigate through our environment without colliding with objects. Equally important in navigation is the proprioceptive information optical flow provides about our own motion (Gibson, 1979), and the motion of objects and other people (Johansson, 1971). Motion is also important as a stimulus to drive eye movements (Rashbass, 1961), and for the segmentation of images into different physical objects.

**Existing theories of motion perception**

There are two distinct strategies prevalent in current theories of motion perception: the feature-based approach, where edges or corners are detected and tracked over time (e.g. Deriche & Faugeras, 1989), and intensity-based approaches which compute motion from the output of spatiotemporal filters without passing through a feature extraction stage. This thesis details the extension of two intensity-based approaches to incorporate adaptive temporal filters. Different intensity-based approaches have in common the use of spatiotemporal filters operating on the image to produce an estimate of velocity. They differ in the form of their filters and the way in which the filter outputs are combined.
**Intensity-based approaches**

Intensity-based approaches can be sub-divided into four classes: gradient, energy, phase and correlational schemes (Figure 1.1). The first three rely on the fact that motion can be considered as orientation in space-time (Figure 1.2) to exploit the analogy with orientation in static 2-D images. Gradient-based approaches use filters which take "fuzzy derivatives" (Koenderink & van Doorn, 1987), blurring and differentiating the image, and combine the filter outputs as a quotient of temporal and spatial derivatives to estimate velocity (Fennema & Thompson, 1979; Horn & Schunk, 1981; Johnston, McOwan & Buxton, 1992; Johnston & Clifford, 1995ab; Young & Lesperance 1993). Energy- and phase-based approaches calculate velocity from the representation of signal structure provided by quadrature pairs of bandpass filters (Gabor, 1946). Each quadrature pair can be thought of as a complex bandpass filter, so the response of each pair to a real image is also complex. Energy-based approaches calculate velocity from the amplitude of this complex signal (Adelson & Bergen, 1985; Heeger, 1987), phase-based approaches from the phase (Fleet, 1990; Fleet & Jepson, 1990). Evidence for a quadrature phase relationship between adjacent cells in the visual cortex has been reported by Pollen & Ronner (1981). Correlation-based approaches compute motion according to a delay-and-compare principle (Reichardt, 1961; van Santen & Sperling, 1985). Under the correlational approach, the basic unit of motion detection consists of a pair of spatially displaced receptors, with the motion signal being computed as the product of the delayed output of one unit with the output of the other.
Figure 1.1: Schematic diagram of strategies for the detection of motion. In the correlation model, motion is detected by filters separated in space and time. The motion energy model requires the construction of oriented filters in space-time. In both the correlational and the Reichardt models, the final output is a comparison of the responses of leftward and rightward detectors. The gradient model compares the response of filters tuned to flicker (T) and static pattern (S) by calculating the ratio of the flicker to pattern responses. Under the phase-based approach, the image is filtered initially with space-time oriented filters of the form used in the energy model. Motion is then computed from the spatial and temporal gradients of the phase output of the filters. Thus, while the initial stage of filtering in the phase model is identical to the energy model, the method of motion computation is the same as under a gradient scheme.
Figure 1.2: Illustration of the idea of motion as orientation in space-time. Frames of an image sequence taken at times T1 to T5 show a grey bar moving across a white background at a constant speed. The sequence of frames may be placed together to form an image volume, with time as the third dimension. A slice through the space-time volume illustrates the fact that motion is equivalent to a space-time orientation. In the case of continuous motion, the bar traces out a shape which is sheared due to motion. Space-time plots are used in chapter 2 to depict spatially one-dimensional motion stimuli.
It is interesting to consider the relationships between the four classes of intensity-based approach. The Elaborated Reichardt Detectors proposed by van Santen & Sperling (1985) involve spatial pre-filtering to avoid problems of aliasing. If the spatial displacement and temporal delay between the filters are both set to be equivalent to a phase shift of $\pi/2$, then van Santen and Sperling (1985) showed that their Elaborated Reichardt Detector is mathematically equivalent to the motion energy detector of Adelson & Bergen (1985), although the different implementations give rise to different physiological predictions regarding the responses of constituent cells (Emerson, Bergen & Adelson, 1992). In certain simple cases, it has also been shown that energy and gradient schemes are mathematically equivalent (Adelson & Bergen, 1986; Simoncelli & Adelson, 1991). Adelson and Bergen demonstrated that a least squares gradient estimator of velocity based on Gaussian derivative filters can be redescribed as an opponent energy model based on filters oriented diagonally in space-time. Adelson and Bergen's derivation relies on the fact that the derivative of a spatiotemporal filter kernel in any given direction can be written as a linear sum of the kernel's spatial and temporal derivatives. Using a 2-D Gaussian kernel, the derivative in any direction will be oriented in space-time (Freeman & Adelson, 1991). However, if a filter with an asymmetric temporal profile is used then the derivative in a given space-time direction will not generally be oriented along a single, dominant direction. Thus it is not generally the case that the spatial and temporal
derivatives of a filter kernel can be linearly combined to produce a
spatiotemporally oriented filter suitable for motion energy calculation. In
particular, the equivalence does not hold for asymmetric temporal filter profiles
necessary in the design of causal filters. Thus, while both energy and gradient
models can employ causal filters, the two classes of model will not be equivalent
in such cases. Since the problem of motion perception demands temporal
causality, any physically realisable gradient scheme will not be equivalent to a
motion energy approach based upon spatiotemporally oriented filters. There is
clearly also a relationship between gradient schemes, based on derivatives of low-
pass filtered images, and phase-based schemes using derivatives of the phase
output of band-pass filtered images. However, we lack as yet a wholly
satisfactory understanding of phase-based techniques, and consequently the exact
nature of the relationship is not yet apparent.

**Methodological issues in the study of biological vision**

Compared to visual psychophysics and electrophysiology, the emergence of
computational methods for studying visual perception has been relatively recent.
While the experimental techniques have been refined over the course of more than
a century, the attempt to characterise biological vision in computational terms
(which I shall term "computational vision") only began to develop in the 1970's
with the application of information processing principles. Foremost among the
theorists during these early years was David Marr, who attempted to give
computational vision a philosophical and representational framework upon which
to build (Marr, 1982). Marr saw the visual system as a complex information processor, recognising that such a system could be understood on a number of levels, but stressed the critical importance of the level of computational theory in understanding both the problems to be solved in visual perception and the solutions adopted by biological systems.

Given the historical development of the study of vision, it is interesting to consider what methodologies might be relevant for today's work. While we have a wealth of empirical data, a strong theoretical basis from which to understand it has emerged only recently through the computational approach. Given this lag of theory behind experimental work, it is unreasonable to expect current theories to explain all the available data, or even to attempt to do so. But this leaves us with the problem of how to evaluate theories that are incomplete or limited in scope. Moreover, if there is to be profitable interaction between experimental and computational vision, we must ask what is the proper role of empirical work in the evaluation of emerging theories.

In order to address these issues, the next section introduces Lakatos' notion of a "research programme" and discusses his ideas on theoretical and empirical progress. These methodological points are illustrated by the characterisation of the computational study of vision as a Lakatosian research programme and the discussion of what might constitute progress within the field, issues regarding the formation of theories, and the relationship between theories and models. The methodological position taken is then advocated for a proposed extension of the
correlational approach to motion perception to include adaptive temporal filters.

*Computational vision as a Lakatosian research programme*

The previous section discussed the way in which the historical development of the study of biological vision has determined which methodological issues are most relevant today. I have chosen to characterise the computational study of vision as a Lakatosian research program because, in trying to explicate how science is a rational enterprise, Lakatos treats science very much as a historical process (Lakatos, 1970, 1978). Lakatos sees science as a succession of theories, in which the new theories developed often preserve many of the important characteristics of the theories they are replacing. According to Lakatos, successive and contemporaneous theories are linked together into a research programme by a "hard core" of common assumptions shared by all investigators. Around this core is a "protective belt" of auxiliary assumptions changeable to accommodate new evidence. Progress within a research programme consists of developing new theories in the protective belt while leaving the core assumptions unchallenged. Lakatos makes an important distinction between theoretical progress and empirical progress. Theoretical progress involves extending the scope of a theory by applying it to new empirical domains, while empirical progress consists of the experimental corroboration of new claims made in the course of theoretical progress. For a research programme to be counted as progressive it must be making both theoretical and empirical progress.
Attempting to characterise computational vision as a research programme in the Lakatosian framework shows the necessary interdependence of visual psychophysics and electrophysiology on the one hand and computational modelling on the other. The empirical disciplines are necessary to generate data and test hypotheses. Computational models serve as illustrations of theories, forcing them to be explicitly specified, and to generate predictions to be tested empirically.

The basic assumptions underlying computational vision have been discussed in depth by Ullman (Ullman, 1991). These include the assumption that the brain is principally an information processing device, that its computations can meaningfully be studied on their own, that the visual system is sufficiently modular for different modalities such as the motion sub-system to be studied separately, and that, at least at the lower levels, processing is primarily bottom-up. While justifiable to a greater or lesser extent by neurobiological or psychophysical evidence, these are all still assumptions. These basic assumptions form part of the conventional wisdom within computational vision. Any actual theories which have not yet been received into the conventional wisdom will necessarily contain auxiliary assumptions, for example that there are specialised mechanisms to deal with different classes of stimuli, or that perceptual mechanisms adapt according to their input.

In order to make theoretical progress in computational vision, the scope of a theory must be extended to apply to new domains. These new domains can either
be within the same modality, for instance applying a theory of motion perception to a further class of motion stimuli, across modalities by drawing upon an analogy such as that between the recovery of image transformations over time necessary for motion perception and those between left and right images for depth in stereoscopic vision, or across species to show convergent evolution. However, to make empirical progress requires experimental confirmation of the claims made on the basis of theory. It is here that computational vision depends crucially on experimental studies if it is to have an impact on the study of biological vision. For without empirical confirmation it cannot hope to be a progressive research programme. In order to be viable, computational vision must rely on visual psychophysics and electrophysiology to provide an empirical test of computational theories.

**Adaptive motion analysis in biological vision**

In previous studies of human vision, adaptive and matched filters have been proposed to account for the effects of subthreshold summation in static images (Hauske, 1974; Hauske, Wolf & Lupp, 1976) and the form of the spatiotemporal contrast sensitivity function (Rovamo, Luntinen & Näsänen, 1993; Watson, Barlow & Robson, 1983). The methodology used here for studying motion adaptation in biological vision involves the interaction of computational modelling and psychophysical and physiological experiments. An adaptive correlation-based scheme, based on the well-established Reichardt detector (Reichardt, 1961; Egelhaaf et al, 1989), is developed to account for the rapid
neural adaptation to motion observed in insect, mammalian and human visual systems. The adaptive correlator is used to model physiological data recorded from the insect lobula plate (figure 1.3) in chapter 2, and the mammalian nucleus of the optic tract (figure 1.4) in chapter 3. The lobula plate neurons are believed to process motion information involved in the stabilisation of flight (e.g. Shi Jian & Horridge, 1989), while the NOT forms part of the accessory optic system, which plays a role in the control of eye movements (Simpson, 1984). Psychophysical evidence of rapid neural adaptation to motion in humans is presented in chapter 4.

Adaptive motion analysis in machine vision

Adaptive filters have found application in many fields for tasks such as system identification, control, and interference cancelling (Widrow & Stearns, 1985; Clarkson, 1993). Here, they are applied to the problem of image velocity measurement. In a comparative study of optical flow techniques (Barron et al, 1994), the most reliable estimates of image velocity were found to be produced by phase-based methods (Fleet & Jepson, 1990). The accurate measurement of image velocity can be highly computationally intensive, particularly in the case of phase-based methods which require the implementation of a number of complex band-pass spatiotemporal filters. In chapter 5, the phase-based approach is described in detail, and an application of adaptive temporal filters is proposed to reduce the number of filters required, and hence the computational overhead, for a given accuracy of image velocity measurement.
Figure 1.3: Schematic diagram of the position in the fly's visual system of the horizontal cells of the lobula plate. The lobula plate receives, through the lamina and the medulla, retinotopic projections from the photoreceptors. In the lobula plate, local movement is spatially integrated by the giant horizontal cells. (Redrawn from Egelhaaf & Borst, 1989, figure 1).
**Figure 1.4:** Neural pathways to the nucleus of the optic tract (NOT); the brain is viewed from below. Pathways via the left hemisphere have been omitted, but are symmetrical to those shown. Solid lines represent pathways carrying information from the right eye, dotted lines from the left eye, and dashed lines connections originating in the cortex. Each NOT is marked with an arrow showing the direction of motion to which it responds. (Redrawn from Braddick, 1996, figure 1.)
Under the adaptive scheme proposed in chapter 5, the temporal filters used to implement the phase-based scheme adapt on the basis of a model of local image translation. Following Horn and Schunk (1981), nearly all current models of motion perception, including phase-based approaches, assume that image velocity is locally constant over a small region of space-time, such that image motion in that region is purely translational. If motion in an image is translational then, in the frequency domain, all the power in the image lies on a plane through the origin (figure 1.5). The translational motion of a spatial profile corresponds to a shearing of its frequency spectrum parallel to the temporal frequency axis, with the angle of shear related to the velocity of the translation (Watson & Turano, 1994). Under the proposed scheme, the filters adapt their temporal frequency tunings such that they move towards the plane of power in frequency space corresponding to the measured image velocity in the region of the image on which they operate. In this way, the filters tend to maximise the signal power in their output.

*Steering or shearing?*

The adaptive scheme proposed in chapter 5 is based on a model whereby image motion corresponds to a shearing of space-time and, hence, a shear of the frequency domain. This is in contrast to previous applications of adaptive filtering to motion analysis, which have operated on the basis of spatiotemporal and frequency domain rotations (Freeman & Adelson, 1991; Haglund, 1992).
Figure 1.5: In the frequency domain, under translational motion, all the power in an image lies on a plane (in 1-D: line) through the origin. (a) Space-time plot of a 1-D image sequence in which a set of black stripes move to the left over time at constant velocity. (b) Doubling the spatial frequency of the striped pattern but keeping the velocity (and hence the orientation in space-time) the same doubles the temporal frequency. (c) Increasing the velocity of the striped pattern without changing its spatial frequency increases the temporal frequency. (d) Plot of temporal versus spatial frequency of the striped patterns in each of the three sequences. All the power in the first two sequences, where motion is at the same velocity, lies on a single line through the origin. The faster velocity in the third sequence gives a point in the frequency domain lying on a more steeply angled line through the origin.
The latter approaches rely on "steering" the weights attached to the outputs of a fixed set of basis filters to interpolate the response of a filter of the same form as the basis set but of arbitrary spatiotemporal orientation. Steering is thus a form of multidimensional adaptation, while the shearing approach taken here is one-dimensional, involving temporal adaptation but fixed spatial filters. While steerable filters are applicable to domains where the dimensions are of the same form (e.g., purely spatial filtering), they are not ideal for problems involving a temporal dimension. In contrast, the shearing approach forwarded here employs causal filters suitable for temporal adaptation.

Recursive filters

As mentioned in the previous section, temporal filters must be causal in that they can only rely on past information, whereas spatial filters can integrate symmetrically over an image. Consequently, while the computational theory underlying the detection of orientation is common to both domains, there will necessarily be differences in the implementation of spatial and temporal filters. A natural implementation for causal temporal filters which reflects the dynamic nature of temporal filtering is provided by recursive (IIR) filters (Bozic, 1986; Papoulis, 1984). Recursive temporal filters have been used for implementations of phase-based (Langley & Fleet, 1992) and gradient schemes (Fleet & Langley, 1995), improving efficiency and reducing the delay inherent in FIR temporal filters. Most current models make the simplifying assumption of using non-causal, non-recursive filters. However, it seems appropriate to view temporal
filters as dynamic entities through recursive implementation when making the filters adaptive rather than fixed in their temporal response properties (Widrow & Stearns, 1985; Clarkson, 1993), and so the adaptive scheme of chapter 5 is implemented using recursive adaptive temporal filters.

Overview

The perception of motion is of fundamental importance in vision, and motion information is required for many higher-level visual processes. There are several current theories of motion perception, and many well-documented biological phenomena to be explained. However, as Holland (1992) commented when writing about the central nervous system, the adaptive systems framework gives us a fresh look at familiar facts, occasionally suggesting new consequences which might otherwise be overwhelmed in the plethora of behavioural and physiological data. This thesis treats the neural and computational substrate of vision as an adaptive system, and investigates the incorporation of adaptive temporal filters into existing motion algorithms. The correlation-based Reichardt detector is extended to include adaptive temporal filters in order to account for existing physiological data from insects (chapter 2). The model is refined to account for physiological data from a mammalian visual system (chapter 3), and the predictions of the model regarding rapid adaptation to motion in humans are tested and confirmed psychophysically (chapter 4). Adaptive signal processing techniques also have potential application in artificial vision: in chapter 5 an adaptive phase-based scheme is developed with a view to accurate and robust
estimation of velocity in noisy image sequences. The relationships between the various classes of intensity-based scheme imply that an extension of any of the existing theories to incorporate temporal adaptation should generalise straightforwardly to other classes of motion algorithm.
Chapter 2

An adaptive Reichardt detector model of temporal adaptation in fly motion vision
In this chapter the applicability of ideas from adaptive signal processing to a correlation-based motion detection scheme are demonstrated. It was work on insect vision that led Reichardt (1961) to propose the first correlation-based motion detector. The precise retinotopic mapping of the fly visual field onto the optic ganglia and the accessibility of its visual system to physiological and neuroanatomical investigation make the fly particularly suitable for the study of the basic principles underlying motion processing (Zaagman et al, 1983; Strausfeld, 1989). Here, a computational model is proposed to account for the adaptive properties of the fly motion system, and the functional benefits of such a scheme are discussed in terms of the coding and processing of image information. The response properties of motion-sensitive neurons in the fly are modelled using an underdamped adaptive scheme to adjust the delays of temporal filters in an array of Reichardt detectors. It is shown that the increase in both temporal resolution and sensitivity to velocity change observed following adaptation to constant motion can be understood as a consequence of local adaptation of the filter delays on the basis of the outputs of elementary motion detectors.

**The Reichardt detector**

Each Reichardt detector consists of two subunits tuned to opposite directions of motion (Figure 2.1). The output of each subunit is analogous to the correlation of two spatially and temporally displaced samples of the image. The response of the motion detector is the difference of the outputs of its two subunits, the sign of the detector response indicating the direction of motion. The peak temporal frequency tuning of a Reichardt detector is inversely related to the delay of its temporal filters. This becomes clear if we consider the response of a Reichardt detector to a moving sine-wave grating. At all times, two signals from the image
reach the correlation stage of the detector, one from each receptor, but one of them has been delayed by a temporal filter. The peak response occurs when signals from corresponding parts of the image reach the correlation stage at the same time. This happens when the time taken for a particular phase of the grating to move between the two receptors is equal to the delay of the detectors' temporal filters.

It is generally accepted that some form of Reichardt detector underlies fly motion detection. The Reichardt detector has also been shown to account for aspects of mammalian motion detection (Ibbotson et al, 1994), and has been put forward in elaborated form as a mechanism of human motion detection (van Santen & Sperling, 1985; Borst & Egelhaaf, 1989; Zanker, 1995). This suggests that strategies employed by insect visual systems may be of general relevance, a theme which will be explored further in chapters 3 and 4. Of specific interest here is the stimulus dependent behaviour observed in the fly motion analysis system, suggesting the operation of adaptive filters.
Figure 2.1: Each Reichardt detector consists of two subunits tuned to opposite directions of motion. The output of each subunit may be thought of as analogous to the correlation of two spatially and temporally displaced samples of the image. The response of the motion detector is the difference of the outputs of its two subunits, the sign of the detector response indicating the direction of motion.
Adaptation in fly motion analysis

This section reviews the experimental evidence for adaptation in fly motion analysis, and examines existing ideas on the site of adaptation within the visual system, the information on which it is based, and its functional significance.

The phenomenology of adaptation

Several electrophysiological investigations into the dynamic response properties of fly motion detectors report adaptation of H1 (Maddess & Laughlin, 1985; de Ruyter van Steveninck et al, 1986; Borst & Egelhaaf, 1987). H1 is a wide-field, directionally-selective, motion-sensitive interneuron located in the lobula plate, the posterior part of the third optic ganglion. De Ruyter van Steveninck et al (1986) find that the characteristic decay time of the velocity step response (Srinivasan, 1983) of H1 depends on stimulus history. By adapting and testing first in the same part of the receptive field of an H1 neuron, then in different regions, they find that the value of the time constant depends on local velocity, from which they conclude that both the determination of the value of the time constant and the actual filtering with the time constant are operations that take place locally within the visual field. Despite the directional selectivity of the neuron itself, de Ruyter van Steveninck et al (1986) find that the time constant depends on the magnitude of the velocity and not its sign. However, Maddess & Laughlin (1985) and Borst & Egelhaaf (1987) report adaptation not to velocity per se but rather to stimulus temporal frequency, the latter pointing out that such an adaptive scheme retains the benefits of an increased operating range and faithful representation of fast transient motion enjoyed by the velocity-adaptive scheme. Borst & Egelhaaf (1987) elicit adaptation
by temporally modulating the luminance of a spatially uniform field, a stimulus containing no velocity information at all. Maddess & Laughlin (1985), while reporting only negligible adaptation to a contrast-reversing grating, find that adaptation to translating stimuli is strongly dependent on temporal frequency. Adaptation has also been reported in the responses of another class of wide-field movement sensitive neuron, the horizontal (HS) cells (Borst & Egelhaaf, 1987; Egelhaaf & Borst, 1989).

The origin of the adapting signal

The dominant view over the last few decades has been that motion detection in the fly visual system can be described by a multiplicative interaction of signals coming from adjacent visual elements (Egelhaaf et al, 1989). Under such a correlation scheme, the filter time constant is reflected in the time course of the decay of the motion detector signal when a moving pattern which stops abruptly is used as a stimulus (de Ruyter van Steveninck et al, 1986). Since wide-field motion-sensitive neurons such as H1 are generally assumed to receive input from a large array of elementary motion detectors (Borst & Egelhaaf, 1987), adaptation of the time constant of the velocity step response of H1 is consistent with adaptation of the filter time constant. While de Ruyter van Steveninck et al (1986) and Borst & Egelhaaf (1987) concur on the site of adaptation, there is disagreement over the origin of the adapting signal. De Ruyter van Steveninck et al (1986) propose that the fly visual system uses estimates of local stimulus velocity to tune its filtering operations, whereas Borst & Egelhaaf (1987) suggest that the adaptational state of a motion detector is governed mainly by the temporal frequency of the signal in its input channels. In the model proposed here, the adapting signal originates
from the locally integrated response of a 1-D array of elementary motion detectors, and is fed back to adapt the time constants of their temporal filters.

*Establishing the functional significance of adaptation*

To establish the functional significance of adaptation it is necessary to consider the response properties of H1 in its adapted state. During the course of adaptation to constant motion the response of H1 declines to a level which shall be referred to as the adapted steady-state response, as distinct from the other time-varying component of the response which shall be termed the transient response. The idea of a steady-state response which changes over time deserves further explanation. In the case of a non-adaptive filter, the onset of a stimulus will give rise to a transient response which will decay over time to leave only the steady-state response of the filter which, as its name implies, remains unchanged as long as the stimulus is maintained. However, in the case of an adaptive filter the steady-state component of the response will change over time as the filter adapts, reaching the adapted steady-state response as adaptation is completed. Here it is the adapted steady-state response which is of interest. In chapter 3 the scope of the model is broadened to encompass transient properties as well.
Figure 2.2 Schematic diagram of adaptive Reichardt detector model. The locally integrated response of a 1-D array of Reichardt detectors is fed back to adapt the delay filters.
The adaptive Reichardt detector model

A novel computational model is proposed to account for the temporal response properties of the H1 neuron of the fly under adaptation. The model is presented in the form of a correlational motion detector (Reichardt, 1961) with adaptable delays, shown schematically in Figure 2.2. The locally integrated responses of an array of elementary motion detectors are fed back to adapt the delay filters prior to correlation. In this way adaptation to a constant motion stimulus reduces the duration of the delay and so tends to null out the response to the maintained motion stimulus. However, the adaptive scheme is underdamped to ensure stability. This underdamping leads to imperfect nulling and hence a non-zero adapted steady-state response to constant motion. Both the duration of the adapted delay and the level of the adapted steady-state response in the model vary as functions of the temporal frequency of the adapting stimulus in qualitatively the same manner as found empirically (Maddess & Laughlin, 1985). Moreover, the behaviour of the model in its adapted state to increments or decrements of velocity around the adapting value is in good agreement with observations of "velocity contrast" measurement (Maddess & Laughlin, 1985; Shi Jian & Horridge, 1991), as defined below. The temporal response profile of an elaborated version of the model is also compared to measurements of the membrane potential of HS neurons in response to motion stimuli (Egelhaaf & Borst, 1989).
Figure 2.3 Amplitude spectra of spatially integrated Reichardt detector arrays at each of two peak tunings for (left) pure delay filters and (right) first-order low-pass filters. In each case the solid line corresponds to the amplitude spectrum in the unadapted state and the dotted line corresponds to the amplitude spectrum after adaptation to a stimulus with temporal frequency of 0.05 units on the horizontal scale shown. The amplitude spectra at temporal frequencies less than that at which the peak response occurs are seen to be largely unaffected by the form of filter.
In modelling the adaptive properties of H1 it is steady-state properties which are of primary interest. In this chapter the delay filters are simulated with a pure delay to eliminate temporal transients from the model's response and reveal its adaptive behaviour. In chapter 3 a more biologically realistic low-pass filter is used to model the impulse response of the delay filter (Egelhaaf & Reichardt, 1987) and thus predict the overall form of the response of H1 more closely. At temporal frequencies less than that at which the peak response occurs the steady-state responses are very similar. Figure 2.3 shows the amplitude spectra of spatially integrated Reichardt detector arrays at each of two peak tunings, incorporating either pure delay or first-order low-pass filters. The magnitude of the response to a moving sine grating is a function of the spatial and temporal phase shifts between samples of the image, and the degree of intensity modulation. For pure delay filters, the response, \( R \), of a 1-D array of movement detectors is given by:

\[
R = (\Delta I)^2 \sin(k\xi) \sin(\omega \tau)
\]  

(2.1)

where \( \xi \) and \( \tau \) are the spatial and temporal offsets between samples, and \( k \), \( \omega \) and \( \Delta I \) are the spatial and temporal frequencies of the grating and its amplitude (see Appendix 2.1). The sine grating has proven a particularly popular stimulus in investigations of the fly visual system, and has been used to obtain all of the data modelled below (Maddess & Laughlin, 1985; Shi Jian & Horridge, 1991; Egelhaaf & Borst, 1989). Here the adaptive behaviour of temporal
filters is of interest, so it is convenient to manipulate the stimulus temporal frequency while keeping the other parameters constant. Under these conditions equation 2.1 reduces to:

\[ R \propto \sin(\omega \tau) \]  

(2.2)

where both the temporal delay, \( \tau \), and the response, \( R \), can vary.

Spatial integration

H1 is known to be a wide-field neuron, assumed to integrate the responses of a 1-D array of movement detectors. It has been shown that the determination of filter time constants takes place locally within the visual field (Maddess & Laughlin, 1985; de Ruyter van Steveninck et al, 1986). It cannot, therefore, be the case that adaptation of the time constants is based simply on the response of H1, since H1 integrates motion information globally over its receptive field. However, to remove modulations present in the responses of elementary movement detectors prior to the calculation of the adapting signal, a degree of local spatial integration is introduced into the model (see Appendix 2.1). Thus spatial integration is modelled as a two-stage process, the first stage involving local integration of elementary movement detector responses, and the second the integration of the first stage’s outputs by H1. Adaptation in the model is based on the responses of units implementing the first stage of spatial integration, post-synaptic to the elementary movement detectors but pre-synaptic to H1. In chapter 3 this assumption is relaxed, and the behaviour of the model is investigated
when low-pass temporal filters are adapted on the basis of the responses of individual Reichardt detectors.

It should be noted that, in all the experiments simulated in this chapter, the stimuli used were spatially uniform. In such cases the responses of all units carrying out the first stage of integration are identical to each other and to the result of subsequent spatial averaging carried out by H1. Thus the model implemented does not involve a second stage of spatial averaging, and the signal, \( R \), which controls adaptation is identical to the response of H1.

**Adaptation to flickering stimuli**

The response of a Reichardt detector is a function of stimulus temporal frequency, but is produced by a directionally opponent combination of sub-units. This is consistent with the observation that counterphase flicker is a poor adapting stimulus, indicating that the basis of adaptation is postsynaptic to the computation of motion, while the strength of adaptation to a moving stimulus is governed by its temporal frequency (Maddess & Laughlin, 1985). To account for the fact that adaptation does occur to some degree in response to temporally modulated stationary patterns (Borst & Egelhaaf, 1987) we would have to assume that some temporal frequency information is available postsynaptic to the movement detectors. A possible source of such information might be imperfect cancellation at the directionally opponent stage due to asymmetries in the movement detectors (Egelhaaf et al, 1989). This idea is explored quantitatively in chapter 3.
Adaptive dynamics

In equation 2.2 we saw that at any instant the response of an array of correlational motion detectors, and hence of the model, depends on stimulus temporal frequency and the delay of the temporal filters. However, since the filter delay is an adaptable parameter, we have two unknowns in equation 2.2, \( R \) and \( t \). Thus the interdependence of \( R \) and \( t \) over time in the adapting system must also be specified. A simple differential equation is proposed to describe the adaptation of the filter delay over time as a function of the locally integrated motion detector response. The absolute response is fed back to reduce the duration of the delay such that temporal resolution increases, and the response desaturates, under adaptation to moving stimuli. The magnitude of the response is employed to ensure adaptation to motion in either the preferred or non-preferred direction, as reported with H1 (de Ruyter van Steveninck et al., 1986; Borst & Egelhaaf, 1987). To prevent the duration of the delay tending to zero a term is included which determines the stable value of the delay in the absence of stimulation. The adaptive dynamics of the model are described by the following differential equation:

\[
\frac{d\tau(t)}{dt} = -\eta \tau(t) |R(t)| + \mu(\tau_0 - \tau(t)) \tag{2.3}
\]

where \( \tau(t) \) is the filter delay, \( \tau_0 \) the unadapted delay, \( R(t) \) the response, \( \eta \) the adaptation rate, and \( \mu \) a parameter controlling the amount of leakage or damping. Equation 2.3 tells us that the filter delay tends to decrease with increasing response magnitude at a rate proportional to the existing delay, but that this decrease is moderated by a tendency for the delay to remain
close to its unadapted value. The response properties of the model are discussed in detail in the next section, when the results of computer simulations are presented and compared with previous experimental results. The theoretical basis of the adaptive scheme in terms of signal coding and transmission is then addressed in the context of existing theories.

Response properties

The temporal response properties of the model are compared to those of the H1 neuron of the fly under adaptation. Both the duration of the adapted delay and the level of the adapted steady-state response of the model vary as functions of the temporal frequency of the adapting stimulus in qualitatively the same manner as found empirically (Maddess & Laughlin, 1985), and the behaviour of the model in its adapted state to increments or decrements of velocity around the adapting value is seen to be in good agreement with observations of velocity contrast measurement (Maddess & Laughlin, 1985; Shi Jian & Horridge, 1991). The temporal response profile of an elaborated version of the model is also compared to measurements of the membrane potential of HS neurons (Egelhaaf & Borst, 1989).
Figure 2.4 (Top) Space-time plot of the onset, maintenance and offset of a sine grating stimulus moving at constant velocity. (Left) The recorded spike rate of H1 to such a stimulus (redrawn from Maddess & Laughlin, 1985, figure 3a). (Right) The response of the model with an unadapted delay, $\tau_u$, of 3.00 frames, an adaptation rate, $\eta$, of 0.05, and a damping parameter, $\mu$, of 0.015, as defined in equation 3. In both the recorded and the model data, the response rises rapidly to a peak after the onset of motion and then decays more slowly to a steady level. In simulations where there is no adaptation the response also rises rapidly but then remains at the peak level without decaying.
Steady-state properties

The response of the model tends asymptotically towards an adapted steady-state in response to a maintained motion stimulus, as illustrated in figure 2.4. In the adapted steady-state the model has ceased to adapt and its tuning depends on the adapting stimulus. Solving equation 2.3 in the steady-state case, \( d\tau/dt = 0 \), gives:

\[
|R_a| = \frac{\mu}{\eta} \left( \frac{\tau_0}{\tau_a} - 1 \right)
\]

(2.4)

where \( R_a \) and \( \tau_a \) are the adapted steady-state response and the duration of the adapted delay respectively (see Appendix 2.2). Equation 2.4 describes the relationship between the magnitude of the adapted response and the adapted delay. However, the independent variable in adaptation experiments is the temporal frequency of the stimulus. We are thus interested in the magnitude of the adapted response and the adapted delay as a function of the temporal frequency of the adapting stimulus. In the absence of temporal aliasing, the magnitude of the response from a Reichardt detector with a given delay increases with stimulus temporal frequency (see figure 2.3). If the model is adapted to a particular constant motion stimulus then increasing the temporal frequency of that stimulus will increase the response of the model, and the delay will have to be reduced in order to return to a steady-state. Thus the adapted delay is a monotonically decreasing function of the temporal frequency of the adapting stimulus. Together with equation 2.4 this indicates that the level of the adapted response will be a monotonically increasing function of temporal frequency. The results of
the model after adaptation to a sine grating translating at constant velocity are shown in figure 2.5. It should be noted that the model predicts the magnitude of the adapted delay to be an almost linear function of temporal frequency when the latter is plotted on a logarithmic axis, and the level of the adapted response to be a compressive function of temporal frequency. These predictions are consistent with the findings of Maddess & Laughlin (1985).
Figure 2.5 (a) Measured time constant (redrawn from Maddess & Laughlin, 1985, figure 9b) (b) duration of adapted delay in model's filters as a function of the temporal frequency of the adapting stimulus. Steady-state response of (c) H1 (redrawn from Maddess & Laughlin, 1985, figure 7) (d) the model as a function of the velocity of the adapting stimulus.
Recovery from adaptation

The dynamics of adaptation in the model are described by the antagonistic interaction of two tendencies, which can be likened to physical forces, described by the two terms on the right-hand side of equation 2.3. The first term represents the force tending to null out the response to sustained motion by reducing the duration of the delay. The second represents the tendency to keep the delay as near as possible to its unadapted length, ensuring stability of the adaptive scheme. The first term is proportional to \(|R|\), so we can think of the strength of the nulling force as being weighted by the response magnitude. When the model is presented with constant motion the delay adapts to a steady value which can be thought of as an equilibrium between the two opposing forces. If the motion ceases the response of the model will drop to zero, as will the strength of the force tending to reduce the duration of the delay. The only force operating after the cessation of motion will tend to return the delay to its unadapted state. The rate at which the delay reverts to this state is given by the second term on the right-hand side of equation 2.3. Solving this differential equation reveals that the delay approaches its unadapted value exponentially over time during recovery (see Appendix 2.3):

\[
\tau(t) = \tau_0 - (\tau_0 - \tau_a) e^{-\mu t}.
\]  

Since the value of the delay increases with time of recovery from adaptation, the response to a subsequently presented stimulus would be expected to show a similar dependency. Maddess & Laughlin (1985) define recovery time as the time elapsed from cessation of the adapting stimulus, and take as a measure of response recovery the peak response to a short burst of
motion. These conventions are followed in figure 2.6, with the peak response to a motion burst plotted against the logarithm of recovery time. The graph shows a monotonic increase of response recovery over time, roughly linear against log time but flattening out at the shortest durations, similar in trend to that found empirically by Maddess & Laughlin (1985). The data of Maddess & Laughlin (1985) show a clear monotonic increase, but the error bars associated with the measured response recovery at each point suggest that the deviation from linearity of the experimentally obtained plot may not be significant.
Figure 2.6 Peak response to a burst of motion as a function of recovery time after the offset of the stimulus for (a) H1 (redrawn from Maddess & Laughlin, 1985, figure 3b) (b) the model.
Adaptation to a sustained motion stimulus causes a drop in the model's response, as shown in figure 2.4. Maddess & Laughlin (1985) suggest that a possible function of adaptation is to desaturate the response of H1 so that it becomes more sensitive to changes in motion around the sustained level. As evidence for this view they examine the responses of H1 during adaptation to a sustained motion. They find that the H1 cell's response to all velocities of the test stimulus is lower after adaptation. They also find that the slope of the response curve in the region of the adapting velocity is steeper after adaptation, demonstrating that H1 becomes more sensitive to changes in velocity around that value. That we should expect similar behaviour from the model is apparent from inspection of its amplitude spectrum at two different delays shown in figure 2.3. Figure 2.7 shows the output of the model to a constant velocity adapting stimulus during which there are short bursts at a faster or slower velocity, revealing the same increase in relative sensitivity and decrease in absolute sensitivity during adaptation.

Interestingly, there is a difference between the responses of H1 and the model to the onset of motion in figure 2.7. The initial response of H1 is sharply peaked while that of the model is rounded. The difference is explained by the presence of a transient component in the response of H1 in addition to the adapting steady-state response being modelled. In chapter 3 the model is equipped with first-order low-pass filters giving rise to a transient component in its response. Consequently, the simulated responses presented in chapter 3 show a sharper peak to motion onset, comparable to those observed in H1.
Figure 2.7 (Top) Space-time plots of sine grating stimuli moving at constant velocity except for regular bursts of (left) faster and (right) slower motion. (Middle) Responses of H1 (redrawn from Maddess & Laughlin, 1985, figure 4), and (Bottom) the model. In all cases, the response to the baseline velocity decreases over time with adaptation while the magnitude of the change in response accompanying increments or decrements of velocity increases.
In analogy with the signalling of photometric contrast by photoreceptors, Maddess & Laughlin (1985) suggest that H1 might be coding for velocity contrast, \( c_v \), defined as follows:

\[
c_v = \frac{\Delta v}{v_m}
\]

(2.6)

where \( v_m \) is the mean velocity and \( \Delta v \) is the change from the mean. It should be noted that velocity contrast is equivalent to temporal frequency contrast, analogously defined, for any constant spatial frequency stimulus (Shi Jian & Horridge, 1991). If H1 is coding for velocity contrast, the change in response produced by a fluctuation in velocity will be proportional to the velocity contrast and independent of the mean velocity. Approximately this behaviour has been found in two separate studies (Maddess & Laughlin, 1985; Shi Jian & Horridge, 1991).

Figure 2.8 shows the output of the model to stimuli of the type used by Shi Jian & Horridge (1991), who modulated stimulus velocity sinusoidally over time around a mean level. They allow the response of H1 to a maintained motion stimulus to reach a steady level and then modulate its velocity sinusoidally in time around the adapting speed. They find that the response of H1 is sinusoidally modulated in phase with the modulation of stimulus velocity, and that the amplitude of the response modulation increases with the degree of velocity modulation. The behaviour of the model agrees qualitatively with the empirical findings of Shi Jian & Horridge (1991). However, when velocity contrast is high, the change in the response of the model to increments of velocity is lower than to decrements, as reported experimentally by Maddess & Laughlin (1985). This finding is attributable to the fact that the elementary motion detectors' response curve flattens out with increasing temporal frequency,
as shown in figure 2.3. Adaptation helps to alleviate this problem by pushing the peak of the response curve to higher temporal frequencies. Figure 2.8 shows that response modulation increases almost linearly with velocity contrast, and is essentially unaffected by variations in mean velocity.
Figure 2.8 (Top Left) Space-time plots of velocity modulated sine grating stimuli at modulation depths of (left) 0.0, (middle) 0.3 and (right) 0.9 at two different velocities (top and bottom). (Top Right) Responses of the model at four different modulation depths. Response modulation of (Bottom Left) H1 (redrawn from Shi Jian & Horridge, 1991, figure 3) (Bottom Right) the model as a function of velocity contrast, each at two background temporal frequencies (■ 2.48 Hz; ● 1.24 Hz. in the original experiment, and arbitrary units in the same ratio for the simulation).
Figure 2.9 shows the response of the model to stimuli with the identical mean velocities and the same velocity contrasts, but in which velocity is modulated at different rates. In each case the response is modulated sinusoidally over time in phase with the velocity modulation, showing that the model is responding to fluctuations in velocity rather than to acceleration. The depth of the response modulation is independent of the rate of velocity modulation. These findings are in agreement with those of Shi Jian & Horridge (1991) for H1.
Figure 2.9  **(Top Left)** Space-time plots of velocity modulated sine grating stimuli with the same mean velocities and velocity contrasts but different modulation frequencies.  **(Top Right)** Response of the model at three different modulation frequencies. Response modulation as a function of modulation frequency for **(Bottom Left)** H1 (redrawn from Shi Jian & Horridge, 1991, figure 4a)  **(Bottom Right)** the model. Modulation frequencies used were 8°/s. (●), 12°/s. (▲), 16°/s. (▼) and 20°/s. (■) in the original experiment, and arbitrary units in the same ratio for the simulation.
Modelling HS

The response of the HS cell is simulated using a modified version of the model of H1, as detailed below. Electrophysiological data from the HS cell (Egelhaaf & Borst, 1989) show the superposition of a transient response and an adapting steady-state response. To model the transient properties of HS closely, temporal integration was introduced both prior to and subsequent to the multiplicative combination of signals in the Reichardt detector. A pure delay filter is retained inside the movement detectors, but additional first-order low-pass pre- and post-filters are introduced. The pre-filter can be thought of as simulating the temporal transfer properties of the photoreceptors, while the post-filter introduces a degree of temporal integration at the spatial integration stage of the original model. Figure 2.10 shows the response of the augmented model to the onset, maintenance and offset of a sine grating moving with constant speed. The model's response captures the essential features of the physiological recordings by Egelhaaf & Borst (1989). The onset of motion gives rise to transient oscillations in the membrane potential of HS at the temporal frequency of the moving stimulus. The amplitude of these transient oscillations decays over time. This transient response is superposed on the steady-state response of the cell, which itself is adapting over time. The cessation of motion produces a transient depolarization of HS, which is found to decay with a time constant dependent on the temporal frequency of the stimulus. To model the adaptation of the offset transient it is assumed that the low-pass post-filter is adapted on the same basis as the movement detector delay.
Figure 2.10 (Top) Space-time plot of the onset, maintenance and offset of a sine grating stimulus moving at constant velocity. (Next to Top) Membrane potential of HS (redrawn from Egelhaaf & Borst, 1989, figure 3). (Next to bottom) Result of Egelhaaf & Borst’s (1989) simulation. (Bottom) Response of the model proposed here.
Coding efficiency

The response properties of the H1 neuron in its adapted state present us with an apparent paradox: H1 seems to be signalling changes in velocity over time, since changes in its response approximate velocity contrast coding; yet the steady-state response of the cell is a monotonic function of the velocity of the adapting stimulus. It is as though H1 is coding both for relative velocity and absolute velocity, and during the course of adaptation absolute sensitivity is being sacrificed for relative sensitivity.

Problems with differential coding

Efficient coding is necessary to reduce the transmission of redundant information over visual pathways of finite bandwidth while maintaining a representation robust to transmission errors. A simple method for increasing the accuracy with which fluctuations in a signal can be transmitted across a channel of finite bandwidth is to subtract a pre-calculated estimate of the signal from the signal itself prior to transmission. Differential or predictive coding was devised to solve a related problem in the field of digital communications (Gibson, 1980), and has also been proposed to occur in biological vision systems (Srinivasan et al, 1982). The effectiveness of differential coding schemes depends on the statistical characteristics of the signal being transmitted, best performance occurring in cases where the distribution of signal values is grouped around a mean level such that small deviations from the mean are more common than large deviations (Barlow, 1972).
Contrast coding is simply a form of differential coding in which the magnitude of fluctuations about the mean level are predicted to be proportional to that level, such that a relative rather than an absolute difference is transmitted. There is, however, a serious problem inherent in differential coding strategies: the effect of transmission errors on their performance (Gibson, 1980; Clarkson, 1993). This problem is illustrated by considering the behaviour of a discrete-time differential encoder which transmits the difference between the current and previous values of a sampled signal. Given knowledge of the initial value of the original signal, the transmitted signal can be decoded and the original signal reconstructed by keeping a running total of the differences transmitted. However if, at any time, there is an error in transmission, the wrong difference value will be received and the values of the original and reconstructed signals will diverge. A robust form of coding is required if convergence of the original and reconstructed signals is to be ensured. Of course, to continue processing a transmitted signal in a biological system it may not be necessary to reconstruct it explicitly, but the same coding considerations apply in ensuring that information is accurately and reliably transmitted (Bruckstein et al, 1983).

"Leaky" predictive coding

Applications of predictive coding in digital communications generally employ an adaptive filter to provide a prediction of the subsequent value of the signal. To ameliorate the effect of transmission errors, the coefficients can be updated according to a "leaky" algorithm (Widrow & Stearns, 1985; Clarkson, 1993). A drawback of introducing leakage is that the operation of the adaptive filter leads to a biased prediction. Under a leaky scheme, a proportion of the original signal is transmitted in addition to the difference information that would be
transmitted following unbiased prediction. The transmission of some redundant information causes the effect of transmission errors to decay over time, at the cost of sub-optimal performance in noiseless situations (Clarkson, 1993).

There is a relationship between the adaptive scheme employed here, described in equation 2.3, and the leaky least mean squares (LMS) algorithm (Widrow & Stearns, 1985; Clarkson, 1993). Under the leaky LMS algorithm a known or estimated signal is adaptively encoded for transmission across a channel of finite bandwidth subject to transmission errors (both limitations of any real communication system). The leaky LMS algorithm can be shown to converge on a solution which minimises mean square error subject to a constraint on the filter coefficients (Darlington & Xu, 1991). Similarly, we can think of the model operating to minimise its response magnitude subject to the constraint that the duration of the delay remains as close as possible to its unadapted value. Analysis of the proposed model of H1 is complicated by the fact that detection and encoding of motion information are achieved at the same stage of computation, the operation of adaptive Reichardt detectors, rather than as successive detection and encoding steps. To examine its behaviour the differential equation in 2.3 can be cast in discrete form as a first difference:

\[
\tau(T+1) - \tau(T) = -\eta \tau(T)|R(T)| + \mu(\tau_0 - \tau(T)) \quad (2.7)
\]

where \( T \) is discrete time, so that the change in the delay at each time step is a function of its existing value and the response magnitude. This is the form of the equation used in computer
simulations of the model. Rearranging gives a leaky update equation for a single coefficient filter (see Appendix 2.4):

\[ \tau'(T+1) = (1-\mu)\tau'(T) + \eta'|R(T)| \]  \hspace{1cm} (2.8)

where \( \tau' = \tau_0 - \tau \), \( 1-\mu \) is a constant leakage factor, \( \eta' = \eta \tau(T) \) is the adaptation rate, and \( R(T) \) can be seen as the (biased) prediction error transmitted. In the case where \( \mu = 0 \), the second term in equation 2.8 will tend to increase the value of \( \tau' \) until \( \tau' = \tau_0 \) (equivalent to a delay of zero duration) causing the response magnitude to fall to zero. Values of \( \mu \) greater than zero prevent the system from going unstable in this way as the value of \( \tau' \) leaks away, damping the adaptation and resulting in convergence to a non-zero value of the delay, \( \tau \).

Adaptation in response to a maintained motion stimulus reduces the magnitude of the model's response, as would be expected under a simple differential coding scheme. However, the response to a maintained stimulus drops not to zero but to a steady level dependent on stimulus parameters, as though a prediction biased to underestimate the response of the model were being subtracted. As detection and encoding of the motion signal are carried out in a single stage, no explicit calculation and subtraction of a prediction takes place. Nevertheless, the adoption of a particular adapted state by the model, defined by the value of \( \tau \), implicitly predicts the form of the subsequent stimulus. We therefore describe the proposed adaptive scheme as a leaky algorithm carrying out biased predictive coding. The rationale for such a model of the behaviour of H1 is the robust and efficient transmission of motion information.
Alternative coding strategies

Zaagman et al (1983) were the first to talk in terms of an adaptive strategy underlying fly motion analysis, concluding that temporal adaptation serves to improve temporal resolution and minimise motion blur. Maddess & Laughlin (1985) observe that the strategy shows some of the characteristics of predictive coding, as outlined by Srinivasan et al (1982), but that it differs in allowing a portion of the mean signal to be transmitted. The strategy being employed here can be understood as a form of leaky or biased predictive coding. The observed similarities between the motion adaptation of H1 and the luminance adaptation of large monopolar cells in the fly retina (Laughlin, 1989; Laughlin, 1990) suggest that this adaptive approach may be of wider relevance than insect motion analysis alone. In the area of human visual perception, a psychophysical study of the detection of changes in motion (Dzhafarov, 1993) found that reaction time to a change in the speed of motion depends on the magnitude of the speed change and is virtually independent of absolute velocity. This would be consistent with some form of differential or contrast coding based upon sensory adaptation, as observed in the fly. However, Dzhafarov et al (1993) have an alternative explanation, which they term "subtractive normalization". Their algorithm treats the detection and encoding of motion changes as distinct, and involves an explicit subtraction of an estimate of the preceding motion. Thus Dzhafarov et al (1993) achieve differential coding in two stages instead of one, without adaptation of the motion detectors. It is conceivable that such a scheme could be proposed to account for the velocity contrast coding observed in the fly, but an additional explanation would then be required for the increased temporal resolution observed following motion adaptation. Alternatively, it may be possible to model some of the adaptive behaviour of the H1 neuron through adaptation of its spike generating mechanism.
(Bruckstein & Zeevi, 1979; Zeevi & Bruckstein, 1981), although again it is not clear that the increased temporal resolution found after motion adaptation could be accounted for in this way.

Summary

A model of adaptive motion analysis in the fly visual system has been presented. The model accounts for some important aspects of electrophysiological data from the H1 neuron, in particular the increase in temporal resolution and the approximation to velocity contrast coding observed following adaptation to motion. It is argued that the function of adaptation is to enable robust, efficient transmission of motion information through a process of leaky predictive coding. Such a computational strategy can be implemented using Reichardt detectors with adaptive time constants. However, the possible realisations of a scheme for leaky predictive coding are not restricted to the Reichardt framework, and it is proposed that analogous strategies may be employed in the coding of visual information in modalities other than motion and in visual systems beyond those of insects. The following two chapters will develop further the ideas put forward here and apply them to motion adaptation in mammalian and human vision.
Appendices

Appendix 2.1

Here we derive the spatially integrated response of an array of Reichardt detectors with pure delay filters to a sine grating stimulus, \( I(x,t) \):

\[
I(x,t) = I_0 + \Delta I \sin(kx - \omega t)
\]  

(2.9)

where \( k \), \( \omega \) and \( \Delta I \) are the spatial and temporal frequencies of the grating and its amplitude, and \( I_0 \) is its mean luminance.

A Reichardt detector with pure delay filters samples the image at four discrete spatiotemporal locations. Letting \( \xi \) and \( \tau \) be the spatial and temporal offsets between samples:

\[
\begin{align*}
I_0 &= I_0 + \Delta I \sin(kx - \omega t) \\
I_1 &= I_0 + \Delta I \sin(kx - \omega (t+\tau)) \\
r_0 &= I_0 + \Delta I \sin(kx+\xi - \omega t)) \\
r_1 &= I_0 + \Delta I \sin(kx+\xi - \omega (t+\tau))
\end{align*}
\]

(2.10)

where \( I_0 \), \( I_1 \), \( r_0 \) and \( r_1 \) are the undelayed and delayed inputs to the left and right branches of the Reichardt detector.

The response, \( S \), of the Reichardt detector is given by the difference of the correlated signals:

\[
S = I_0 r_1 - r_0 I_1 .
\]

(2.11)

Substituting for \( I_0 \), \( I_1 \), \( r_0 \) and \( r_1 \) gives:

\[
S = (\Delta I^2 [\sin(kx - \omega t) \cdot \sin(kx+\xi - \omega (t+\tau)) - \sin(kx+\xi - \omega t) \cdot \sin(kx - \omega (t+\tau))] + I_0 \Delta I [\sin(kx - \omega t) + \sin(kx+\xi - \omega (t+\tau)) - \sin(kx+\xi - \omega t) - \sin(kx - \omega (t+\tau))].
\]

(2.12)

If \( R \) is the mean response of a 1-D spatial array of \( N \) Reichardt detectors with responses \( S_i \):

\[
R = \frac{1}{N} \sum_{i=1}^{N} S_i
\]

(2.13)
then:

\[ R = (\Delta I)^2 \left[ \sin(kx - \omega t) \sin(k(x + \xi) - \omega(t + \tau)) - \sin(k(x + \xi) - \omega t) \cdot \sin(kx - \omega(t + \tau)) \right] \]  

(2.14)

since integrating over space removes the term in \( I_o \Delta I \). Simplifying using trigonometrical identities gives the desired result:

\[ R = (\Delta I)^2 \sin(k\xi) \sin(\omega t) \]  

(2.1)

Appendix 2.2

Here we derive the adapted steady-state response of the model. The model can be said to have reached an adapted steady-state when the duration, \( \tau(t) \), of the Reichardt detectors’ delay filters has ceased to change in response to a maintained stimulus. In this case, \( d\tau/dt = 0 \) and \( \tau(t) \) is a constant, \( \tau_a \).

The dynamic equation governing the adaptation of the delay, \( \tau(t) \), is:

\[ \frac{d\tau(t)}{dt} = -\eta \tau(t) |R(t)| + \mu(\tau_0 - \tau(t)) \]  

(2.3)

where \( R(t) \) is the model response and \( \tau_a, \mu \) and \( \eta \) are constants. Substituting in the adapted steady-state conditions stated above gives:

\[ 0 = -\eta \tau_a |R(t)| + \mu(\tau_0 - \tau_a) \]  

(2.15)

Rearranging in terms of the response magnitude:

\[ |R(t)| = \frac{\mu}{\eta} \left( \frac{\tau_0}{\tau_a} - 1 \right) \]  

(2.16)

we see that the right-hand side of the expression is composed entirely of constants, showing that, for constant stimulus attributes, the model’s adapted steady-state response is a constant, \( R_a \), where:
Here we derive the dynamic equation describing the model's recovery from adaptation. After adaptation to a maintained stimulus the duration of the filter delay tends towards a steady value, $\tau_a$, which depends on the adapting stimulus. If the temporal variation in the stimulus ceases abruptly then the response of the model will also drop to zero (see equation 2.2). Substituting $R(t) = 0$ into equation 2.3 gives the dynamic equation describing the model's recovery from adaptation:

$$\frac{d\tau(t)}{dt} = \mu(\tau_0 - \tau(t)). \tag{2.17}$$

Substituting $\tau'(t) = \tau_0 - \tau(t)$, where $\tau'(t)$ is the difference of the duration of the delay from its unadapted value gives:

$$\frac{d\tau'(t)}{dt} = -\mu \tau'(t) \tag{2.18}$$

since $d\tau(t)/dt = -d\tau'(t)/dt$. Rearranging and integrating gives:

$$\int_{0}^{\tau'} \frac{d\tau'(t)}{\tau'(t)} = -\mu \int_{0}^{t'} dt \tag{2.19}$$

where $t'$ is the time elapsed since the cessation of the adapting stimulus and $\tau_a' = \tau_0 - \tau_a$ (where $\tau_a$ is the adapted steady-state delay). Evaluating the integrals gives:

$$\ln \left( \frac{\tau'(t')}{\tau_a} \right) = -\mu t'. \tag{2.20}$$

Taking the inverse logarithm of both sides and rearranging gives:
\[
\tau'(t') = \tau_a' e^{-\mu t'}.
\]

(2.21)

Substituting for \(\tau'(t')\) and \(\tau_a'\) in terms of \(\tau(t)\) and \(\tau_a\) and rearranging gives:

\[
\tau(t) = \tau_0 - (\tau_0 - \tau_a) e^{-\mu t}
\]

(2.5)

where here \(t\) denotes the time elapsed since the cessation of motion.

**Appendix 2.4**

Here we show that the adaptive dynamics of the model, described in continuous time by equation 2.3, can be approximated in discrete time by a leaky update equation. A discrete form of equation 2.3 can be obtained by approximating the derivative, \(d\tau(t)/dt\), by a first difference, giving:

\[
\tau(T+1) - \tau(T) = -\eta \tau(T)|R(T)| + \mu (\tau_0 - \tau(T))
\]

(2.7)

where \(T\) is discrete time. This is the form of the equation used in computer simulations of the model.

We define, \(\tau'(t)\), as the difference of the filter delay from its unadapted value: \(\tau'(t) = \tau_0 - \tau(t)\). Substituting in equation 2.7 and rearranging gives:

\[
\tau'(T+1) = (1 - \mu) \tau'(T) + \eta \tau(T)|R(T)|.
\]

(2.22)

Defining \(\eta' \tau(T) = \eta \tau(T)\) as an adaptation rate proportional to the current value of \(\tau(t)\) gives:

\[
\tau'(T+1) = (1 - \mu) \tau'(T) + \eta' |R(T)|.
\]

(2.8)

It can be seen that the above equation is of the form of a leaky coefficient update equation for a single coefficient filter:

\[
f(T+1) = \alpha f(T) + \beta e(T)
\]

(2.23)

where \(f(T)\) is the filter coefficient at time \(T\), \(e(T)\) is an estimate of the error between the desired and obtained responses to be minimised subject to the effects of \(\alpha\), a constant
leakage factor (where $\alpha = 1$ represents no leakage), and $\beta$ is the adaptation rate (where $\beta = 0$ represents no adaptation). In this way the model adapts to minimise the magnitude of its response subject to the tendency for $\tau(t)$ to decay to zero (equivalent to $\tau(t)$ returning to its unadapted value, $\tau_0$) as a result of leakage.
Chapter 3

Extending the model to mammalian motion adaptation
Introduction

In the previous chapter we reviewed evidence of the adaptation to motion of wide-field, direction-selective neurons in the fly lobula plate (Maddess & Laughlin, 1985; de Ruyter van Steveninck et al, 1986; Borst & Egelhaaf, 1987). Adaptation of motion-sensitive neurons over the course of a few seconds has also been reported in the optic lobes of the butterfly (Maddess et al, 1991), in area 17 of the cat visual cortex (Vautin and Berkley, 1977; Hammond et al. 1988; Maddess et al, 1988; Giaschi et al, 1993), and in the nucleus of the optic tract (NOT) of the wallaby (Ibbotson & Mark, 1996). Here, physiological data on motion adaptation is recorded from cells found in the NOT of adult wallabies, *Macropus eugenii* (methods as in Ibbotson et al, 1994). Neurons in the NOT provide the signals that control the slow phases of horizontal optokinetic nystagmus (e.g. Collewijn, 1975a,b; Hoffmann & Schoppmann, 1981; Schiff et al, 1988; Simpson, 1984; Soodak & Simpson, 1988). They are directional and respond preferentially to wide-field stimulation.

Marked similarities are revealed between the adaptation to motion observed in wide-field directional neurons found in the mammalian nucleus of the optic tract and cells in the insect lobula plate. It is found here that the velocity impulse response (Srinivasan 1983) in wallaby NOT neurons depends on stimulus history within local regions of the neuron's receptive field, as with cells in the fly lobula plate (Maddess & Laughlin, 1985; de Ruyter van Steveninck et al, 1986). The corresponding temporal frequency response functions are found to be shifted laterally and compressed by motion adaptation. The lateral shift serves to enhance dynamic range and differential motion sensitivity. It will be argued that the compression is not caused by fatigue but is an intrinsic property of the adaptive process, resulting from the interdependence of temporal frequency tuning and gain in the temporal filters of the motion detectors. However, while the form and time scale of adaptation is comparable in the two systems, there is a difference in the directional properties of the effect.
in that wallaby NOT cells appear to adapt much less strongly to anti-preferred motion and flicker than do cells in the fly lobula plate (Borst & Egelhaaf, 1987).

The response to motion of both fly lobula plate neurons (Egelhaaf et al, 1989) and cells in the NOT of the wallaby (Ibbotson et al, 1994) are consistent with the computation of motion by a correlation-based Reichardt detector (Reichardt, 1961). In chapter 2, a model based on the Reichardt detector was proposed to describe motion adaptation in the insect visual system. In this chapter, the model is reformulated to account for the key features of motion adaptation in cells found in the mammalian NOT and the insect lobula plate, with only minor modifications required to account for the observed differences in directionality between the two phyla.

**Model**

In chapter 2, the increase in both temporal resolution and sensitivity to velocity change of neurons in the fly lobula plate following motion adaptation were modelled by adjusting the delay filters in an array of Reichardt detectors. However, there were three points on which the previous model's performance was not satisfactory: (a) only adaptive steady-state behaviour was accounted for, not transient responses; (b) data on directionality and flicker sensitivity (Borst & Egelhaaf, 1987) were not predicted; and (c) a degree of spatial integration of motion detector responses was required prior to obtaining the signal upon which adaptation was based, while experimental evidence shows that such integration could only operate over a very limited range (Maddess & Laughlin, 1985; de Ruyter van Steveninck et al, 1986).

In the previous chapter an idealised version of the Reichardt detector was used, involving pure delay filters and a perfectly balanced subtraction stage in order to eliminate temporal
transients from the model's response and focus on its adaptive behaviour (Fig. 3.1a). The limitations of the model proposed in chapter 2 are largely due to the idealised nature of the detector implemented. Egelhaaf et al (1989) described a version of the Reichardt detector which, while not adaptive, accurately described other aspects of lobula plate physiology. Here we extend the adaptive Reichardt detector model to incorporate the realistic motion detector characteristics described by Egelhaaf et al (1989), and to predict the spatially localised nature of adaptation. Three significant modifications to the model are made which improve its predictive power and increase its scope to account for data from neurons found in the wallaby NOT as well as the fly lobula plate. While the revised model can be used to predict the majority of effects observed in both species, it was necessary to introduce slight differences in the form of the model to account for observed differences in the directionality and flicker-sensitivity of adaptation. These differences are detailed later in the context of results on the directionality of adaptation. The modifications to the model are:

1. First-order low-pass temporal filters (Appendix 3.1) are used in the detectors' input channels (Reichardt, 1961; Egelhaaf et al, 1989). The use of low-pass filters (Fig. 3.1b) enables transient responses to be modelled. In chapter 2 pure delay filters were used (Fig. 3.1a) and, as a consequence, only adapting steady-state behaviour was shown. In order to capture fully the dynamic nature of the adaptive process, a recursive implementation of the temporal filters is now employed (Appendix 3.2).

2. The directionally-opponent subtraction stage is not perfectly balanced (Fig. 3.1b). The subunits in a Reichardt detector do not respond exclusively to motion signals, they also give some motion-independent responses (Egelhaaf et al. 1989). If the subtraction stage is balanced, these motion-independent responses are removed. However, when the subtraction stage is unbalanced, some motion-independent signals are transmitted. As a result, a flickering stimulus induces a response in the
Figure 3.1: Schematic diagrams of (a) the previous version of the model, (b) the model proposed here. The models differ in their temporal filters, the opponency stage, and the origin of the adapting signal.
output of the detector. Small responses to flicker have been observed in NOT neurons and lobula plate neurons (Ibbotson et al. 1994; Egelhaaf et al. 1989), suggesting that the subtraction stages in the real systems are not perfectly balanced.

3. Adaptation of the model presented here is based directly on the responses of single Reichardt detectors (Fig. 3.1b; Appendix 3.1). In the model of chapter 2, the temporal frequency dependent signal that drives adaptation was obtained after a degree of spatial integration (Fig. 3.1a). However, experimental evidence suggests that integration could only operate over a very limited spatial range, considerably smaller than the receptive field size of a wide-field NOT or lobula plate neuron (Maddess & Laughlin, 1985; de Ruyter van Steveninck et al, 1986). A second stage of spatial integration would then be required in order to give a wide receptive field (Fig. 3.1a). Hence the minimum possible locus of adaptation is the size of an elementary motion detector's receptive field.
Results

Temporal frequency dependency of response decay

In response to sustained motion, mammalian NOT neurons and insect lobula plate neurons show similar characteristic properties (Figure 3.2a,b). The response rises rapidly to a peak at motion onset, then decays approximately exponentially to a steady level over the course of a few seconds. The rate of decay of the response increases with temporal frequency, at least at frequencies below the unadapted peak tuning of the motion detectors. This characteristic response behaviour is predicted by the model (Figure 3.2c). When stimulated through a relatively narrow aperture, a distinct response oscillation is superimposed on the mean firing rates of the neurons (Figure 3.3a,b). Such oscillations are also apparent in the model’s output when only a small number of elementary motion detectors are integrated over, simulating the presentation of a spatially restricted stimulus (Figure 3.3c). In both physiological data and model simulation the principal oscillatory component is at the temporal frequency of the stimulus.

Temporal frequency dependency of adapted time constant

If a stimulus grating is jumped by an amount less than or equal to a quarter of a cycle in the preferred direction in a single frame, the cells in the wallaby NOT elicit large impulse responses (Ibbotson & Mark, 1996). Under a correlational motion detection scheme, the time constant of the decay of a wide-field unit’s impulse response reflects the time constants of the temporal filters in the elementary motion detectors feeding into it (de Ruyter van Steveninck et al, 1986). Thus, adaptation of the motion detectors’ temporal filters will lead to a decrease in the time constant of the impulse response decay. Here, impulse responses were recorded from wallaby NOT cells at a range of adapting stimulus temporal frequencies.
Figure 3.2: Response to sustained motion at each of two temporal frequencies from neurons in (a) the wallaby NOT (solid line: 3.04 Hz; dashed line: 18.24 Hz), and (b) the fly lobula plate (solid line: 2.4 Hz; dashed line: 19.2 Hz) (adapted from figure 9a, Maddess & Laughlin, 1985). (c) The response of the model. The wallaby data was obtained using a moving sine wave grating with a spatial frequency of 0.25 cycles/degree.
Figure 3.3: Response of a wallaby NOT neuron to a 0.25 cycles/degree grating moving in the cell's preferred direction with a temporal frequency of (a) 3.04 Hz, (b) 18.24 Hz. (c) The model's response to two different temporal frequencies of a moving sine-wave grating.
There was a 5.5 s adapting phase in which a grating was moved in the preferred direction, followed by a 0.5 s pause in which the screen was blank, followed by presentation of the impulsive stimulus. The impulsive stimulus consisted of the grating appearing in one frame then moving by a quarter cycle in the next frame.

Figure 3.4a shows the impulse response of a neuron adapted with a 0.5 cycles/degree grating moving in the preferred direction with a temporal frequency of 3.2 Hz. The time constants of the exponentials fitted to the impulse responses of this cell were smaller the higher the temporal frequency of the adapting stimulus, up to 0.76 Hz. (Figure 3.4b). At drift rates between 0.76 and 6.04 Hz, which were the frequencies that elicited the largest responses from the cell, the time constants for the decay of the impulse responses remained fairly constant. At frequencies above 6.04 Hz the time constants increased, returning to the unadapted level at 24.32 Hz. The same trends were observed in all 9 cells tested in this way. The time constants derived from the impulse responses of the model also form a U-shaped curve when plotted as a function of adapting temporal frequency (Figure 3.4c), with the smallest time constants occurring at temporal frequencies that match the peak temporal tuning of the model as was the case for the real neurons.

**Directionality of adaptation**

To test the directionality of motion adaptation in the wallaby, several NOT cells were adapted with stimuli moving in the preferred or anti-preferred directions and with gratings of the same spatial frequency that were stationary but reversed their contrast at the temporal frequency of the moving grating. The results from a directional adaptation experiment in one cell are shown in Figure 3.5. The impulse elicited in the control experiment had a slow exponential decline with a time constant of 1.013 s (Figure 3.5a). After a period of adaptation to motion in the preferred direction the impulse response had a shorter time
Figure 3.4: (a) Impulse response of a wallaby NOT neuron elicited by a 0.5° jump of the stimulus in the cell's preferred direction. The decay of the impulse response was fitted with an exponential (solid line). Decay time constants for the impulse responses of (b) the wallaby NOT neuron, (c) the model, as a function of the prior adapting temporal frequency of a moving grating.
Figure 3.5: The impulse response of a wallaby NOT neuron (a) with no prior adaptation; and after 5.5 seconds adaptation to (b) preferred direction motion, (c) anti-preferred motion, and (d) counter-phase flicker.
constant of 0.233 s (Figure 3.5b). When the adapting stimulus was moved in the anti-preferred direction the time constant of the impulse response was 0.847 s (Figure 3.5c), while it was 0.916 s following a period of contrast reversal (Figure 3.5d). The results presented in Figure 3.5 show the general trends that were found in all four neurons tested with these stimuli. Motion in the preferred direction induced the most significant adaptation. Anti-preferred motion and contrast reversal both led to a similar reduction in time constant but the decrease was always less than for adaptation to motion in the preferred direction. These general trends become clear when the time constants for the impulse responses are plotted against the stimulus condition (Fig 3.6a: dotted line). However, in the neurons of the fly lobula plate, significant adaptation is observed in response to motion in either the preferred or the anti-preferred directions and also to flicker (Fig. 3.6a: solid and dashed lines).

Individual Reichardt Detectors elicit highly oscillatory responses when stimulated with moving sine-wave gratings. During motion in the preferred direction the responses are most often above the baseline so the mean output is positive (Fig. 3.6c). The opposite is the case for motion in the anti-preferred direction (Fig. 3.6d). These oscillatory responses at the outputs of the individual motion detectors are attenuated by spatial integration prior to the computation of the wide-field response. To simulate data from the fly lobula plate, the modulus of the difference between the detector response and the baseline was used to provide the adapting signal (Appendix 3.1). That is, the adapting signal was full-wave rectified. Therefore, both preferred and anti-preferred motion produced significant adaptation (Fig. 3.6b: solid line). To simulate the wallaby data, the difference between the detector response and the baseline was used only if the response exceeded the baseline (Appendix 3.1). That is, the adapting signal was half-wave rectified. In this case, preferred motion resulted in much greater adaptation than motion in the anti-preferred direction (Fig. 3.6b: dotted line). In the previous version of the model, presented in chapter 2, the adapting
Figure 3.6: The time constant of the decay of the impulse response after adaptation in four conditions: (C) control (no adaptation), (P) preferred motion, (AP) anti-preferred motion, and (F) counter-phase flicker for (a) H1 and HSE cells of the fly lobula plate (Borst & Egelhaaf, 1987) and the mean of four wallaby NOT neurons, (b) the model with full- and half-wave rectification of the adapting signal. The response of an individual adaptive Reichardt detector to motion in (c) the preferred direction, (d) the anti-preferred direction.
signal was full-wave rectified, resulting in equal adaptation to preferred and anti-preferred motion (Fig. 3.6b: dashed line). Had half-wave rectification been used in that model there would have been no adaptation to motion in the anti-preferred direction.

The unbalanced nature of the subtraction stage in the adaptive Reichardt detector proposed here led to small responses to temporal modulations in the absence of motion. The degree of imbalance in the subtraction stage governed the strength of the flicker response, and hence the degree of flicker adaptation. This feature of the model predicts the adaptation to flickering stimuli observed in both the fly and wallaby data (Fig. 3.6a). It was necessary to unbalance the subtraction stage of the model by about twice as much in the fly as compared to the wallaby to simulate the relative levels of adaptation to flicker and preferred motion. In the earlier model there was no flicker sensitivity and hence no adaptation to flicker.

**Effect of adaptation on temporal frequency tuning**

Adaptation of wide-field motion-sensitive neurons in the insect lobula plate and wallaby NOT causes the peak temporal tuning of the cells to shift to higher frequencies (fly: Maddess & Laughlin, 1985; butterfly: Maddess et al, 1991). Adaptation in these neurons produces not only a shift in the peak temporal tuning, but also a compression in the temporal frequency response function (Fig. 3.7a-c). The peak temporal frequency tuning of a Reichardt detector is inversely related to the time constant of its temporal filters. This is because each temporal filter delays the signal passing through it by an amount proportional to its time constant. Decreasing the time constants, and hence the delays, of the temporal filters, will increase the stimulus temporal frequency necessary to produce the detector's peak response (Fig. 3.7d). The version of the model put forward in chapter 2 also showed a shift in peak temporal frequency tuning following adaptation, but this was not accompanied by compression of the temporal frequency response function (Fig. 3.7e).
Figure 3.7: Temporal frequency response functions in the unadapted state (solid line) and after a period of motion adaptation (dashed line) from neurons in (a) the wallaby NOT, (b) the fly lobula plate (adapted from figure 5a, Maddess & Laughlin, 1985), and (c) the butterfly lobula plate (adapted from figure 9b, Maddess et al, 1991). Temporal frequency response functions before and after adaptation from (d) the present model, (e) the model proposed in chapter 2.
Why does the revised model account for the response compression while the version proposed in chapter 2 does not? In the unadapted state, the temporal filters in the present model have a gain of unity. During adaptation to motion, the time constants of the filters decrease. The filters are not renormalised at any time, so their adapted impulse response initially rises to the same value as in the unadapted state (Fig. 3.8). The shorter time constant causes the adapted impulse response to fall off more rapidly, so the area under the adapted impulse response is smaller than in the unadapted case. The area under the impulse response is a measure of the gain of the system, so adaptation to motion results in a decrease in gain of the Reichardt detectors' temporal filters. This is reflected in a decrease in gain of the wide-field unit which takes its input from the Reichardt detector array, and hence a compression of its temporal frequency response function (Fig. 3.7d), as observed in the real neurons.

Time constant and gain in the model are interdependent because the gain of the temporal filters is not renormalised during adaptation (Appendix 3.3). Renormalisation would require that the initial value of the filters' impulse response be increased as the time constant decreased. In neural terms this would require astronomically high firing rates to motion onset when the filters were in a strongly adapted state. It would seem a more parsimonious use of the cells' response range to use spike rates up to the maximum possible to signal motion in the unadapted state. If this is the case, then decreasing the time constant through adaptation will necessarily reduce gain if maximum firing rates are not to be exceeded.

The earlier version of the model fails to show response compression, as adaptation of the pure delay filters simply rescales the temporal frequency response function along the temporal frequency axis without introducing any other deformations (Fig. 3.7e).
Figure 3.8: The model's impulse response in the unadapted state (solid line) and after a period of adaptation to motion in the preferred direction (dashed line).
Is adaptation generated locally?

It was tested whether adaptation in one isolated region of the receptive field could generate an adaptive effect in another independent region of the field. The stimulus consisted of a circular central region (radius: 7°) surrounded by a circular annulus (inner radius: 8°; outer radius: 12°). The stimulus cycle was the same as that used to investigate the temporal frequency dependency of the adapted time constant. Several independent trials were conducted. First, impulse responses were elicited by an impulsive 0.25 cycle movement of the test stimulus in either the central region or the annulus alone (control). Second, the adapting grating in either the centre or the annulus was moved for 6 s and then the test stimulus was presented in the same stimulus region. Third, the adapting stimulus was presented in the centre and the test in the annulus, or vice versa. Prior adaptation in the same stimulus region led to a reduction in the time constant of the subsequent impulse response (Figure 3.9a). Prior adaptation in the other stimulus region did not lead to any significant change in the time constant of the impulse response. The results demonstrate that adaptation in one region of the receptive field did not influence the response elicited by stimulation in another spatially separated region of the field. The model is able to predict this data as adaptation is generated locally in the model by feeding back the responses of individual motion detectors (Fig 3.9b).

Discussion

Broad similarities in the form and time scale of motion adaptation have been found between neurons in the mammalian NOT and the insect lobula plate. However, there is a marked difference in the directionality of the phenomena. A simple but powerful model of motion adaptation is proposed to account for the observed data. Only minor differences in the form of the model are required to explain the difference in directionality between the two systems,
Figure 3.9: The time constants of the decay of the impulse responses of (a) a wallaby NOT cell, (b) the model, under control conditions, after adaptation to motion in a different region of the receptive field to the test stimulus, and after adaptation and testing in the same region.
namely the degree of unbalancing of the subtraction stage and whether rectification of the
adapting signal is full- or half-wave. The proposed model is a modification of the one
proposed in the previous chapter to account for various phenomena in insect motion
adaptation. In a sense the new model is simpler, in that adaptation is now based on the
response of individual motion detectors prior to spatial averaging. This is perhaps the most
important feature of the model as it explains the local adaptation phenomena described in the
physiological literature (Maddess & Laughlin, 1985; de Ruyter van Steveninck et al, 1986).
Also, the model now employs biologically plausible filters through a recursive
implementation, and does not require mathematically precise subtraction at its opponency
stage. These modifications have increased the scope of the model to include the simulation
of transient responses and the prediction of directional effects.

Adaptation: function or fatigue?

Many studies have investigated aftereffects to motion adaptation, rather than concentrating
on the form and time scale of adaptation as it is actually occurring. Moreover, the duration
of adaptation used is typically of the order of minutes. However, it is imperative to study
adaptation over very short durations in order to gain insights into the functional significance
of adaptation to the behaving animal. Motion adaptation causes a lateral shift in the
temporal frequency response function resulting in an enhanced dynamic range and improved
differential motion sensitivity. The shift in temporal tuning is accompanied by an overall
compression of the response function. Motion adaptation occurs too quickly, and is too
predictable, to be attributable to fatigue. The observed compression of the temporal
frequency tuning curve may look like fatiguing, but it has been shown here that such a
compression can arise simply as a consequence of the adaptation of filter time constants.
Indeed, it may be that motion-sensitive cells avoid fatigue by adapting such that their initial
spike rates are not maintained. This is consistent with the view that neural adaptation is a
functional way of minimising the intrinsic limitations on neural transmission, rather than a
direct consequence of those limitations.

Summary

This chapter has described striking similarities between the adaptive characteristics of wide-field directional neurons in two widely separated phyla, insects and mammals. The main features of adaptation in the NOT are that it occurs locally and independently at loci within the receptive field of each neuron, it is temporal frequency dependent, and it is directional. Adaptation shifts the temporal frequency response functions of NOT neurons along the temporal frequency axis so that high differential sensitivity to changes in temporal frequency are maintained over a wide range of frequencies. It has been shown that an adapting motion detector model, with only slight modifications, can be used to predict the responses in both insects and mammals. Importantly, the response attenuation observed after a period of motion adaptation is shown to be an intrinsic property of the adaptive process and is not caused by fatigue. The following chapter will provide psychophysical evidence that motion processing in humans shares the adaptive characteristics of fly and wallaby motion vision, by showing that a reduction in perceived speed during adaptation to a moving stimulus is accompanied by an increased sensitivity to changes in speed.
Appendices

Appendix 3.1: Time constant adaptation

The impulse response, $h(t)$, of the first-order low-pass temporal filters used in the model is given by:

$$h(t) = \begin{cases} \frac{1}{\tau_0} \exp \left( \frac{t}{\tau(t)} \right) , & t \leq 0 \\ 0 , & t > 0 \end{cases}$$

where $\tau(t)$ is the adaptable filter time constant and $\tau_0$ is its unadapted value. As defined, the filter has a gain of unity when $\tau(t) = \tau_0$. Adaptation of the filter time constant, $\tau(t)$, is governed by the following dynamical equation:

$$\frac{d\tau(t)}{dt} = -\eta \cdot f(R(t)) + \mu (\tau_0 - \tau(t))$$

where $R(t)$ is the time-varying response of the Reichardt detector, $\eta$ is the adaptation rate, and $\mu$ is a damping coefficient. The function $f(R(t))$ determines the properties of the Reichardt detector response that govern adaptation. To simulate fly data, $f(R(t))$ is the modulus of the difference between the detector response, $R(t)$, and the baseline response. Thus, motion in either direction produces strong adaptation. To simulate the wallaby data, $f(R(t))$ is the difference between the detector response and the baseline response only if the response exceeds the baseline. In this case, preferred motion results in stronger adaptation than motion in the opposite direction.
Appendix 3.2: Recursive implementation of first-order low-pass temporal filters

The output, $O(t)$, of any one of the model's temporal filters to an input image sequence can be represented mathematically by the convolution of the filter's impulse response, $h(t)$, with the signal, $I(t)$, received by the receptor:

$$O(t) = \int_{-\infty}^{\infty} h(t). I(t + t') \, dt' = \left[ \frac{1}{\tau_0} \right] \int_{-\infty}^{0} \exp \left( \frac{t'}{\tau(t)} \right) I(t + t') \, dt'$$

where $h(t)$ is the impulse response of a first-order low-pass filter given in Appendix 3.1. To simulate temporal filtering on a digital computer the above continuous-time equation must be cast in discrete form:

$$O(T) = A \cdot \sum_{T'}^{0} \exp \left( \frac{T'}{\tau(T)} \right) I(T + T')$$

where $T$ represents time sampled discretely in frames, and $A$ is the discrete-time normalisation constant for the case where $\tau(T) = \tau_0$.

There are two ways in which the operation of temporal filters in discrete time can be simulated. The continuous-time impulse response can be discretely sampled to produce FIR (finite impulse response) filters, which can then be convolved with our input image sequence. However, the first-order low-pass filters to be simulated have an impulse response which decays exponentially. Thus an FIR implementation would require an infinite number of samples to represent the impulse response without truncation, and would require a similar record of the image sequence to be maintained. The alternative is to implement
recursive IIR (infinite impulse response) filters. Recursive filters maintain an implicit record of the past image sequence by feeding back the response from the previous frame and combining it with new input from the receptors. In this way truncation problems are avoided, and no storage of past images is required. Moreover, the dynamic nature of recursive filters makes them particularly suited to the adaptive scheme, where the filter time constant is subject to change on a frame-to-frame basis. Thus a recursive implementation is used to implement the first-order low-pass temporal filters in the model.

The recursive implementation is derived from the discrete-time equation for the filter response given above. We wish to express the filter output, \( O(T) \), at any given time in terms of the filter output at the previous frame, \( O(T-1) \), and the current input signal from the image, \( I(T) \). Knowing \( O(T) \), we can substitute \( T-1 \) for \( T \) to get \( O(T-1) \):

\[
O(T - 1) = A \cdot \sum_{r=-\infty}^{0} \exp \left( \frac{T'}{\tau(T)} \right) \cdot I(T - 1 + T')
\]

\[
= A \cdot \sum_{r=-\infty}^{1} \exp \left( \frac{T' + 1}{\tau(T)} \right) \cdot I(T + T')
\]

\[
= \exp \left( \frac{1}{\tau(T)} \right) \cdot \left[ A \cdot \sum_{r=-\infty}^{1} \exp \left( \frac{T'}{\tau(T)} \right) \cdot I(T + T') \right]
\]

\[
= \exp \left( \frac{1}{\tau(T)} \right) \cdot [O(T) - A \cdot I(T)]
\]

Rearranging gives us an expression for \( O(T) \) in terms of \( O(T-1) \) and \( I(T) \):

\[
O(T) = \exp \left( -\frac{1}{\tau(T)} \right) \cdot O(T - 1) + A \cdot I(T)
\]

as desired. Normalising for a gain of unity when \( \tau(T) = \tau_0 \) gives:
The resulting recursive implementation is shown diagrammatically in Figure 3.10.

Appendix 3.3: Dependence of peak gain on filter time constant

Adaptation is implemented in the model by updating the value of $\tau(T)$ at every time step, according to a discretised version of the dynamical equation given in Appendix 3.1:

\[
\tau(T) = \tau(T-1) - \eta \cdot f(R(T)) + \mu \cdot (\tau_0 - \tau(T-1)) \\
= (1 - \mu)\tau(T-1) - \eta \cdot f(R(T)) + \mu \cdot \tau_0
\]

Updating $\tau(T)$ alters the weight, $B(T)$, attached to the feedback path of the recursive filter (see Figure 6), since:

\[
B(T) = \exp\left(\frac{-1}{\tau(T)}\right)
\]

The weighting, $A$, of the feedforward path is held constant:

\[
A = 1 - \exp\left(\frac{-1}{\tau_0}\right)
\]

The gain of the filter, $G(T)$, is a function of $A$ and $B(T)$. Since $B(T)$ is a function of the filter time constant, $\tau(T)$, so is the gain:

\[
G(\tau) = \frac{A}{1 - B(\tau)}
\]

The filter time constant, $\tau(T)$, is always positive and finite, so $B(T)$ is always positive but less than one. As the filter time constant, $\tau(T)$, decreases, so the weight attached to the feedback path, $B(T)$, and hence the gain, $G(T)$, decrease also. Thus, as the time constant of the model's filters decreases through adaptation, so their gain decreases also. Since the
Figure 3.10: Schematic diagram of the recursive implementation of the first-order low-pass filter used in the model. The output, \( O(T) \), of the filter is delayed by one frame and fed back to combine with the image signal, \( I(T) \). The weight, \( B(T) \), attached to the feedback signal within the filter depends adaptively on the response, \( R(T) \), of the adaptive Reichardt detector of which the filter is a part. The weight, \( A \), attached to the image signal remains constant.
wide-field response of the output of the model is simply a sum over the responses of individual detectors, a decrease in the gain of the detector responses is reflected by a corresponding decrease in the gain of the response of the wide-field unit. The filter time constant could be decreased neurally by feeding back the output of each elementary motion detector such that it inhibits signal transmission along the feedback path of the filter (Figure 3.10).
Chapter 4

Psychophysics of human motion adaptation parallels insect and mammalian electrophysiology
Introduction

The previous two chapters have discussed rapid neural adaptation to motion in insect and mammalian vision. The response of fly lobula plate (Maddess & Laughlin, 1985) and wallaby NOT neurons to a maintained motion stimulus decreases over time while sensitivity to variations in speed around the maintained level increases through a lateral shift in the temporal frequency response function. This behaviour is modelled by modifying a correlation-based motion detector to include adaptable temporal filters. In this chapter the form and time course of human motion adaptation is investigated psychophysically and compared to the electrophysiological data. If a human observer were making psychophysical judgements based on the response of a population of cells with adaptive characteristics similar to those of neurons in the fly lobula plate or wallaby NOT, then adaptation to a maintained motion stimulus would have two measurable effects:

1. A decrease in perceived stimulus speed with adaptation duration;
2. An increase in sensitivity to variations in speed with adaptation duration.

Two psychophysical experiments are forwarded that support these predictions. First, the effect of adaptation duration on absolute sensitivity to motion is investigated. Subjects are required to perform a speed discrimination between a stimulus maintained at a constant velocity and a short burst of a comparison stimulus. Second, the effect of adaptation duration on sensitivity to modulation of velocity around the
adapting level is established. Subjects are required to detect a brief sinusoidal modulation of the velocity of a maintained stimulus. The perceived speed of a maintained motion stimulus is found to decrease as a function of adaptation duration. Sensitivity to modulation of the velocity of a maintained motion stimulus is found to increase as a function of adaptation duration. The form and time scale of the sensitivity changes are comparable to those observed electrophysiologically in insect and mammalian vision. The results suggest that mechanisms in the human visual system underlying adaptation to motion are comparable to those found in the fly and the wallaby. It is proposed that, in all cases, adaptation serves to improve the transmission of novel motion information along the visual pathways at the expense of maintaining an accurate representation of the unchanging components of the stimulus. This form of adaptation has previously been observed psychophysically in the contrast domain by Greenlee and Heitger (1988), and in orientation by Regan and Beverley (1985). Greenlee and Heitger showed that adaptation to a grating of contrast 0.8 not only increases detection thresholds for gratings of similar spatial frequencies, but also decreases contrast discrimination thresholds for contrast levels above 0.5. Regan and Beverley showed that orientation discrimination for test gratings parallel to a high contrast adapting grating is improved, while detection is degraded.

Method

To measure the effect of perceived speed on adaptation duration, subjects were required to carry out a two alternative forced choice (2AFC) discrimination between
the speeds of a maintained sinusoidal grating and short bursts of a comparison
stimulus (see Figure 4.1a). The maintained grating had a contrast of 25%, a spatial
frequency of 0.75 cycles per degree, and a speed of 8.33 degrees per second. The
contrast of the comparison stimulus was ramped on and off to a maximum value of
25%, its spatial frequency was also 0.75 cycles per degree, and its velocity was
determined under computer control according to the method of constant stimuli.

Stimuli were displayed on a linear Joyce DM 4 monitor with a mean luminance of 8.8
cd/m² using a VSG2/3 graphics display card. Subjects CC and KL were the author of
this thesis and his supervisor, subject CB was an experienced psychophysical observer
naive to the purposes of the study. All subjects had normal vision.

Subjects were required to fixate a central spot from a viewing distance of 126
centimetres. The adapting stimulus appeared in one quadrant of the visual field, in a
circular aperture of 3.3 degrees diameter and 2.0 degrees eccentricity. The adapting
stimulus remained present throughout the 12 second duration of the trial. Inter-trial
adaptation was avoided by ensuring that the interval between presentations of stimuli
of the same orientation in any given quadrant of the visual field was at least five
times the duration for which the adapting stimulus was present (see Figure 4.1a).
Comparison stimuli were presented for 640 milliseconds at intervals of 2.56 seconds.
The onset asynchrony between the adapting stimulus and the first comparison
stimulus was varied randomly between 0 and 1.28 seconds, allowing sampling of the
temporal axis every 1.28 seconds. Psychometric functions were fitted to each
subject's data independently at each adaptation duration tested, each point of
2 AFC Speed Discrimination
Adapting stimulus present for 12 s.
Comparison stimuli presented every 2.56 s.
Each comparison stimulus present for 0.64 s.

Yes-No Detection of Velocity Modulation
Adapting stimulus present for 9 s.
Decision every 1.28 s.
Duration of velocity modulation 80 ms.

Figure 4.1: Schematic description of the experimental procedure used to measure (a) perceived speed, and (b) sensitivity to velocity modulation, as a function of adaptation duration.
subjective equality being used as a measure of perceived speed. It was assumed that
adaptation is spatially localised to the quadrant of the visual field in which the
adapting stimulus is presented, and that effects of comparison stimuli on the
perception of subsequent comparison stimuli within a trial are negligible. In this way,
plots of perceived speed against adaptation duration were obtained for each observer.

To measure sensitivity to velocity modulation as a function of adaptation, a yes-no
detection task was used. In pilot studies with a 2AFC paradigm subjects had reported
an increase in the salience of velocity modulations over time but also increasing
difficulty in localising those modulations. In the yes-no task subjects again had to
fixate a central spot while an adapting stimulus appeared in one quadrant of the visual
field (see Figure 4.1b). Every 1.28 seconds a beep from the computer prompted the
subject to respond whether or not a modulation in the velocity of the stimulus had
been seen. At each beep there was a 50% probability of a sinusoidal modulation of
velocity of 80 milliseconds duration. The number of correct responses at any given
adaptation duration was taken as a measure of sensitivity to velocity modulation at
that duration. Stimulus parameters were the same as for the adapting stimulus in the
previous experiment.

Results

Perceived speed during adaptation

The first hypothesis to be tested is that perceived stimulus speed decreases with
adaptation duration. Figure 4.2 shows the perceived speed of a translating grating stimulus as a function of adaptation duration. For all three observers, perceived speed decreases over time to a level around 75-80% of its unadapted value. For each observer an exponential curve was found to fit the data well, with a half-life of the order of 3 seconds for a stimulus speed of 8.33 degrees per second. This compares with half-lives of up to 2 seconds in the decay of the response of fly H1 neurons (Maddess & Laughlin, 1985) and wallaby NOT cells for faster moving stimuli. The adaptation of both perceived speed and the cell responses are well described by an exponential decay to a steady level.

In order to attribute the reduction in perceived speed of a grating stimulus over time to motion adaptation per se, it is necessary to control for adaptation in perceived contrast. Hammett et al (1994) show that the perceived contrast of a sinusoidal grating fades as a function of time. Moreover, stimulus contrast has an effect on perceived speed such that at low contrasts perceived speed is reduced (Thompson, 1982; Muller & Greenlee, 1994). Taken together, these observations suggest that the adaptation in perceived speed demonstrated here need not necessarily be due to adaptation of elementary motion detector time constants but could simply be an artefact of contrast adaptation.

In order to demonstrate adaptation to stimulus motion per se, the effect of contrast adaptation was controlled for in the following experiment. The previous experiment was repeated but, instead of varying the speed of the test stimulus and holding its contrast constant, test stimulus contrast was varied and its speed held constant.
Figure 4.2: Plot of the perceived speed of an adapting grating as a function of adaptation duration for each of three observers. Observers were required to discriminate the speed of a 0.75 cycles/degree, 6.25 Hz adapting grating and a briefly presented test grating at a range of adaptation durations. Psychometric functions were fitted to the observers' responses independently at each duration, the points of subjective equality being used as measures of perceived speed. Exponential curves were then fitted to describe the change in perceived speed for each observer as a function of time.
Subjects were required to report which stimulus, adapting or test, had the higher contrast. In this way measurements of perceived contrast were obtained as a function of adaptation duration (Figure 4.3ab). These measurements were then used to provide test stimuli perceptually matched in contrast for a further run of the speed matching task. Figure 4.3cd compares the results of the speed matching task using fixed and perceptually matched contrasts. For both observers the form and time scale of adaptation does not differ significantly between the two cases, demonstrating that adaptation of perceived speed occurs independent of contrast adaptation.

Two further experiments were carried out to control for any possible interaction between the motions of the test and adapting stimuli, and to ensure that adaptation was not occurring between trials but only within trials. In the first experiment the perceived speed data was broken down according to the relative directions of the test and adapting stimuli. Were there significant interactions across space between the percepts of motion in the two stimuli, one would expect quantitative differences in the results obtained for motion in the same and opposite directions. However, the results for two observers (Figure 4.4ab) show no such differences. In the second experiment, data collected in the first and second halves of each session were compared to test whether exposure to previous trials had an effect on perceived speed. The results for two observers (Figure 4.4cd) reveal no such effect, showing that inter-trial adaptation was successfully eliminated by the experimental design.

A final perceived speed experiment was carried out to investigate whether it is temporal frequency or velocity which determines the time scale of adaptation. Fly
Figure 4.3: (a,b) Plots of the perceived contrast of an adapting grating as a function of adaptation duration for each of two observers. (c,d) Plots of the perceived speed of an adapting grating using test stimuli of perceptually matched contrasts for each of two observers compared with the fixed contrast case.
Figure 4.4: Perceived speed of an adapting grating as a function of adaptation duration for each of two observers, comparing (a,b) cases in which the adapting and test stimuli were moving in the same and opposite directions, (c,d) data collected in the first and second halves of each trial.
H1 neurons and wallaby NOT cells show their maximum response at a given temporal frequency, virtually independent of stimulus spatial frequency (fly: Eckert, 1980; wallaby: Ibbotson et al, 1994). Such behaviour is consistent with the existence of elementary motion detectors of the Reichardt form (Reichardt, 1961). It has also been shown that the adaptation of the response of fly H1 neurons is governed by stimulus temporal frequency (Maddess & Laughlin, 1985). Given the parallel between the adaptation of perceived speed shown above and the decay of fly lobula plate cell responses, it is predicted that psychophysical observers' perceived speed judgements will also be governed by temporal frequency. To investigate this prediction the perceived speed task was repeated over a range of spatial and temporal frequencies. Results are shown for two observers at three spatial frequencies where temporal frequency (Figure 4.5ab) and velocity (Figure 4.5cd) is held constant. The data is not conclusive in establishing which stimulus attributes determine the time scale of motion adaptation. However, the results shown in Figure 4.5 do demonstrate that a decrease in perceived speed as a function of adaptation duration occurs over a range of spatial and temporal frequencies.

Differential velocity sensitivity during adaptation

The second hypothesis to be tested is that sensitivity to variations in speed increases with adaptation duration. Figure 4.6 shows the sensitivity to velocity modulation as a function of adaptation duration for each of the three observers. For each observer, data are plotted for two different depths of modulation, where modulation depth is defined as the peak magnitude of the velocity change to be detected divided by the
Figure 4.5: Perceived speed of adapting gratings with (a,b) temporal frequencies of 6.25 Hz, (c,d) velocities of 3.125 degrees/second, as a function of adaptation duration at each of three spatial frequencies.
Figure 4.6: Plots of the number of correct responses on a yes-no detection of a sinusoidal modulation of velocity as a function of adaptation duration. Data at two different velocity modulation depths are presented for each of three observers. In each case, the adapting grating had a contrast of 0.25, a spatial frequency of 0.75 cycles/degree, and a speed of 8.33 degrees/second.
baseline velocity of the adapting stimulus. In all cases, sensitivity to modulations around the baseline velocity is found to increase with adaptation duration, in accord with the second hypothesis. These results are in qualitative agreement with the findings from insect (Maddess & Laughlin, 1985) and mammalian vision (chapter 3) that adaptation increases relative sensitivity to motion. This is clear evidence that rapid adaptation to motion occurs in humans, and that it has a functional benefit in terms of the detection of changes in motion.

Discussion

The responses to motion both of a psychophysical observer and of motion-sensitive neurons in the fly and the wallaby are affected by stimulus history. Psychophysically, both perceived speed and the gain of the optokinetic nystagmus (Maddess & Ibbotson, 1992) decay over time in response to a maintained stimulus. Electrophysiologically measured spike rates have been shown to behave in the same way. With the existing data, one cannot be sure whether the difference in the time course of adaptation reflects differences in the temporal properties of the respective visual systems, or is simply due to the difference in stimulus parameters used.

Motion adaptation also results in increased sensitivity to small perturbations in speed. This point is relevant to the debate on whether the human visual system's adaptation to motion has a functional basis (Wright & Johnston, 1985; Verstraten et al, 1994) or is simply a consequence of neural fatigue (e.g. Mather, 1980; Smith & Edgar, 1994). The latter studies involve adaptation durations of at least a minute prior to any form
of testing, and so an element of neural fatiguing cannot be discounted. Here, motion adaptation is shown to occur over the course of a few seconds, and to result in a reduction in perceived speed and an increase in differential motion sensitivity. The rapidity of adaptation, and its functional role in the detection of changes in motion, support the conclusion that it is not neural fatigue that underlies the perceptual changes reported here. The model presented in the previous chapter demonstrated a possible mechanism, without any element of fatigue, by which a reduction in the response to an adapting stimulus coupled with an increase in the slope of the temporal frequency response function around the adapted temporal frequency might be achieved.

Summary

There are clear parallels in the form and time scale of motion adaptation observed psychophysically in humans and electrophysiologically in mammals and insects. It is likely that in all cases the function of adaptation is the efficient transmission of motion information. This information must be represented by the activity of neurons having only a limited number of response states. Best use can be made of these response states by adaptively reallocating them to match the prevailing distribution of stimulus attributes. In this way the effective operating range of the system can be increased. The form of adaptation reported here in the motion domain is also apparent in the domain of spatial contrast in humans (Greenlee & Heitger, 1988) and insects (Laughlin, 1989), suggesting that adaptive coding strategies have wide applicability in biological vision systems over a range of visual dimensions.
Chapter 5

An application of adaptive filters to the measurement of image velocity in machine vision
This chapter proposes an application of adaptive IIR filters to the problem of image velocity estimation. A phase-based motion algorithm (Fleet & Jepson, 1989; Langley & Fleet, 1992) is employed to measure velocity locally within an image sequence from the outputs of a set of complex space-time separable band-pass filters. The filters' temporal tunings are adaptively modified on the basis of measured velocity to optimise the representation of image motion. In computer simulations, the scheme is shown to provide accurate estimates of velocity even at high levels of image noise.

Introduction

The method used here for measuring image velocity is based on the phase information in the output of band-pass filters (Fleet & Jepson, 1989). In a comparative study of the quantitative accuracy of several optical flow techniques (Barron et al, 1994) it was found that phase-based methods produced the most reliable estimates of image velocity. However, a common criticism of the many techniques that have been developed for computing image motion, and one which applies particularly to phase-based approaches, is the computational expense and the amount of memory necessary to store the required number of frames and the different filters' outputs. The phase-based approach of Fleet and Jepson (1989), for example, involved 22 filters at each spatial scale, each of which was an FIR filter with a temporal extent of 15 frames or more.

In an attempt to create motion algorithms that run at or near frame-rates it is
necessary to address the issue of efficiency. In general, the most common way of reducing storage requirements has been to restrict the duration of temporal support to a small number of frames (e.g. 2 or 3). An alternative is to introduce recursive temporal filtering, allowing performance to be maintained while improving markedly on the efficiency of implementation (Langley & Fleet, 1992; Fleet & Langley, 1995). Here, the use of a class of adaptive IIR temporal filters within a phase-based framework is described.

Recursive centre-frequency adaptive band-pass filters were originally designed for the tracking and enhancement of one-dimensional signals in biomedicine and communications (Kumar & Pal, 1985). Kumar and Pal's filters adapt according to a gradient algorithm operating on the basis of power output, while those employed here use measurements derived from phase. Phase information is reasonably stable with respect to small signal perturbations (Fleet & Jepson, 1993), and it is therefore hoped that phase-based adaptation may offer faster convergence and wide applicability in signal processing.

The purpose of introducing adaptation into image velocity estimation is twofold: to reduce computation and storage requirements by reducing the number of independent temporal filters required; and to increase accuracy by maximising filtered image power. Assuming a model of constant image translation, \( I(x,t) = I_r(x-vt) \), image power will be constrained to lie on a plane through the origin in frequency space:
\[ \tilde{I}(k,\omega) = \tilde{I}_s(k)\delta(v.k - \omega), \quad (5.1) \]

where \( k = (k_1, k_2) \) and \( \omega \) denote spatial and temporal frequencies respectively.

Therefore, if \( I_s(x) \) is spatially band-pass, then the spatiotemporal signal \( I(x,t) \) is temporally band-pass, with its temporal frequencies constrained by its spatial frequency content and the image velocity. This situation is very suitable for the use of centre-frequency adaptive temporal filters. Using a model of local image translation, and filters of fixed spatial tuning, measured velocity is fed back to adjust the peak temporal frequency tuning of the IIR filters to match the temporal characteristics of the spatially filtered image signal and so optimise the representation of image motion (see Figure 5.1).

**Measuring velocity from band-pass filters**

The phase-based approach to velocity measurement operates on the outputs of complex band-pass filters and their derivatives, so causal band-pass temporal filters and corresponding derivative filters are required. A family of band-pass filters are used to decompose the input \( I(x,t) \) according to local scale, speed and orientation. Each filter kernel is complex-valued and may be written as:

\[ \Psi(x,t) = W(x,t) \exp [j(k.x + \omega.t)], \quad (5.2) \]
Figure 5.1: Schematic diagram of the method of image velocity measurement described in this chapter. Velocity measurements are computed from the phase outputs of space-time separable band-pass filters and fed back to adapt the filters’ temporal tunings so that they match the space-time characteristics of the image signal.
where $W(x,t)$ is a real-valued window function, and $\exp[jkx] = \cos(kx) + j\sin(kx)$ with $j^2 = -1$. The convolution of these filters with the image sequence is given by:

$$R(x,t) = I(x,t) \ast \Psi(x,t) .$$  \hspace{1cm} (5.3)

Because $R(x,t)$ is complex-valued it may be written as:

$$R(x,t) = \rho(x,t) \exp[j\phi(x,t)] ,$$  \hspace{1cm} (5.4)

where $\rho(x,t)$ and $\phi(x,t)$ are its amplitude and phase components.

Assuming the conservation of the logarithm of the band-pass filtered image signal, $R(x,t)$, i.e. $d[\ln R(x,t)]/dt = 0$ (Fleet, 1992), image velocity can be measured with first-order differential constraints on the amplitude and phase component of $R(x,t)$, as in:

$$[(\frac{\rho_x}{\rho} + j\phi_x)u + (\frac{\rho_y}{\rho} + j\phi_y)v + (\frac{\rho_t}{\rho} + j\phi_t)] = 0 ,$$  \hspace{1cm} (5.5)

where $v(x,t) = (u,v)$ is the local image velocity and subscripts denote partial differentiation. This yields two equations, from the real and imaginary parts:
\[
\phi_x(x,t) u + \phi_y(x,t) v + \phi_z(x,t) = 0 ,
\]

(5.6)

\[
\rho_x(x,t) u + \rho_y(x,t) v + \rho_z(x,t) = 0 .
\]

(5.7)

Here, the phase derivatives of the filter outputs are combined over a small spatiotemporal region to arrive at an estimate of local velocity and a measure of confidence in that estimate. In principle, given image translation, both terms might be used to compute image velocity. The preference for phase over amplitude is a result of robustness considerations (Fleet & Jepson, 1989; Fleet, 1992). Phase is relatively stable under deviations from image translation that commonly occur with projections of 3-D scenes, while the amplitude of the filter response is not. Amplitude is not conserved, for example, under dilation of the image as a camera approaches an object. However, phase has also been shown to exhibit occasional instability, most often in the neighbourhoods about phase singularities. For this reason it is desirable to impose a stability constraint (Jepson & Fleet, 1990; Fleet, 1992) to detect unreliable estimates of velocity. Here, a "signal/certainty philosophy" (Knutsson & Westin, 1993) is followed such that for each local estimate of velocity at any given time there is an associated measure of the confidence of that velocity estimate. The calculation of these confidence measures is discussed below.

To avoid explicit calculation and differentiation of phase, and the attendant problem of phase unwrapping, the following identity is used to extract phase
derivatives (Fleet & Jepson, 1989; Franks 1969):

\[
\phi_i(x,t) = \frac{\text{Im}[R^*(x,t) \cdot R_i(x,t)]}{|R(x,t)|^2}.
\]  

(5.8)

Velocity is computed using a weighted least squares fit of the local first-order constraints on phase to a model of constant velocity in each small spatiotemporal neighbourhood, \( N \), by minimizing:

\[
\sum_{(x,t) \in N} W(x,t) [\nabla \phi(x,t) \cdot v(x,t) - \phi_i(x,t)]^2,
\]  

(5.9)

where \( W(x,t) \) is a window that gives more weight to constraints near the centre of the neighbourhood. The minimization of equation 5.9 leads to a linear system of the form \( W\Phi \cdot v = W\phi_i \), the solution of which is given by:

\[
v = (\Phi_x^T W\Phi_x)^{-1} \Phi_x^T W\phi_i,
\]  

(5.10)

where \( v \) is the image velocity vector, \( \Phi_x \) and \( \Phi_t \) are the \( nx2 \) matrix and \( nx1 \) vector of spatial and temporal phase derivatives, and \( W \) is the diagonal weight matrix.
Calculating confidence estimates

As well as producing reliable measurements of image velocity, it is desirable for an image motion analysis scheme to indicate image regions where the fitting of a local translational velocity model is poor. This is particularly important if the output of the motion analysis scheme is to be combined with data from other low-level visual modules (e.g. form or stereo) to reconstruct the 3-D visual environment and segment it into objects. Here, unreliable estimates of velocity are identified using the eigenvalues of the spatial covariance matrix, $\Phi_x^T W \Phi_x$ (Barron et al, 1994; Fleet & Langley, 1995). The magnitude of the smallest eigenvalue is used as a measure of confidence in the associated velocity estimate. This value depends upon the range of orientations and the magnitude of spatial gradients present locally within the image.

Spatial filters

The spatial filters implemented here are complex band-pass Gabor filters (Gabor, 1946) and their derivatives. The real and imaginary parts of the filter form a Hilbert transform pair, and are thus said to be in quadrature (Figure 5.2). The filters are tuned to each of 6 spatial orientations, with centre-frequency spatial tunings of 0.2 cycles per pixel and envelope standard deviations of 2.5 pixels. Gabor filters are used because of their favourable resolution in the signal and frequency domains (Wilson & Granlund, 1984). The width of the Gaussian envelope was chosen to ensure that the DC response of the real part of the filter is
Figure 5.2: Plot of the form of the spatial filters used to implement the phase-based scheme for velocity measurement. The complex Gabor kernel is plotted over space as its real (solid line) and imaginary (dotted line) parts. In the computer simulations, the filters have a centre-frequency tuning of 0.2 cycles per pixel and an envelope standard deviation of 2.5 pixels.
negligible (Langley, 1990).

**Space-time oriented filters**

The phase-based scheme requires spatiotemporally band-pass filters oriented in space-time. Such filters are separable into one-dimensional spatial and temporal filters. To achieve quadrature the real and imaginary parts of the constituent spatial and temporal filters are combined according to trigonometrical identities in the same way as for energy models (e.g. Adelson & Bergen, 1985). A schematic diagram of the construction of complex space-time oriented linear filters is shown in Figure 5.3.

**Causal recursive temporal filters**

The above phase-based scheme is implemented using space-time separable filters \( \Psi(x,t) = A(t)B(x) \), allowing temporal filtering to be considered in isolation:

\[
R(x,t) = A(t) \cdot [B(x) \cdot I(x,t)] . \tag{5.11}
\]

The temporal filters are designed in the continuous-time domain and then transformed to obtain a discrete-time transfer function. The class of filters used are derived from the truncated exponential, \( E(t) \), a causal low-pass filter with a
Figure 5.3: Schematic diagram of the construction of spatiotemporally oriented filters from pairs of separable filters. The four spatiotemporal impulse responses shown across the top are products of two spatial and two temporal impulse responses (see figures 5.2 and 5.4). The ones across the bottom are sums and differences of those above. The result is a pair of leftward- and a pair of rightward-selective filters. Members of a pair are approximately in quadrature.
nearly linear phase spectrum (Fleet & Langley, 1995):

\[
E(t) = H(t) b \cdot \exp(-bt) = \begin{cases} 
  b \cdot \exp[-bt] , & t \geq 0 \\
  0 , & t < 0 
\end{cases},
\]  
(5.12)

where \(H(t)\) is the Heaviside step function, and \(b^\prime \) is the time constant, a measure of the duration of temporal support. Band-pass filters are obtained by modulating with a complex sinusoid to move the peak tuning frequency away from the origin:

\[
A_1(t) = \begin{cases} 
  b \cdot \exp[-bt + j\omega_0 t] , & t \geq 0 \\
  0 , & t < 0 
\end{cases}.
\]  
(5.13)

The amplitude and phase spectra are given by:

\[
|A_1(\omega)| = \frac{b}{\sqrt{b^2 + (\omega - \omega_0)^2}}, \quad \text{arg}[A_1(\omega)] = \tan^{-1}\left(\frac{\omega - \omega_0}{b}\right).
\]  
(5.14)

The amplitude spectrum has the same shape as that of the low-pass filter (for which \(\omega_0 = 0\)), but is centred at the modulation frequency \(\omega_0\). The phase spectrum is now predominantly linear in the passband about \(\omega_0\), and thus suitable
for use in a phase-based scheme.

*Cascaded temporal filters*

A concern with the truncated exponential is the slow decay of its amplitude spectrum, making localisation of its passband in frequency space poor. However, localisation in the frequency domain can be improved by cascading truncated exponentials. The Central Limit Theorem tells us that, as the number of cascades increases, the resulting filter approaches a Gaussian (Bracewell, 1986), or, if modulated, a Gabor function (Gabor, 1946). Gabor filters have been widely used in motion analysis and stereopsis, in part because of their favourable resolution in the signal and frequency domains (Wilson & Granlund, 1984). \(n\)-fold convolution of \(A_i(t)\) is equivalent to convolution with a single kernel, \(A_n(t)\):

\[
A_n(t) = \begin{cases} 
\frac{t^{n-1} b^n}{(n-1)!} \cdot \exp[-bt + j\omega_0 t], & t \geq 0 \\
0, & t < 0
\end{cases} . \quad (5.15)
\]

From the Convolution Theorem (Bracewell, 1986), the Fourier transform of \(A_n(t)\) is \((\tilde{A}_i(\omega))^n\), the amplitude and phase spectra of which are:

\[
|\tilde{A}_n(\omega)| = \frac{b^n}{(b^2 + (\omega - \omega_0)^2)^{n/2}} , \quad \arg [\tilde{A}_n(\omega)] = n \tan^{-1} \left( \frac{\omega - \omega_0}{b} \right). \quad (5.16)
\]
The amplitude spectrum retains its peak tuning at $\omega_0$, and the phase spectrum of $A_\alpha(t)$ is still reasonably linear. The temporal support and bandwidth of the cascaded filter are determined by a combination of $b$ and $n$. The peak of the impulse response function shifts away from $t = 0$ as $n$ increases, representing an implicit delay in the measurements. It can be shown, by differentiating the envelope, that the peak occurs at time $t = (n-1)/b$. In the simulations described below, $n = 3$ and $b = 0.8$, corresponding to a delay of the order of two and a half frames. The real and imaginary parts of a third-order filter are shown in figure 5.4.

**Digital filter design**

Digital realisation of the phase-based scheme requires discrete IIR implementations of the continuous filter described above (equation 5.15) and its temporal derivative. To generate the appropriate digital filters, a bilinear transform is used to map the Laplace transform ($s$-domain) representation of the filter onto the discrete-time $z$-domain:

$$H(z) = H_c(s)\bigg|_{s \frac{2}{T}\left(1 - z^{-1}\right)}$$

(5.17)

where $T$ is the temporal sampling interval. This method has previously been used
Figure 5.4: Plot of the form of the temporal filters used to implement the phase-based scheme for velocity measurement. The complex impulse response of the third-order filter is plotted over time as its real (solid line) and imaginary (dotted line) parts. In the computer simulations, the filter envelope reaches a peak after 2.5 frames and the modulation frequency is under adaptive control.
for the implementation of low-pass temporal filters (Fleet & Langley, 1995).

Although the impulse invariant method is somewhat simpler, it is susceptible to severe aliasing problems (Bozic, 1986). The bilinear transform avoids aliasing, effectively by warping frequency space (Bozic, 1986), and it is then a simple matter to unwarp the phase derivatives in converting back to the frequency domain.

Using the fact that convolution in the time domain is equivalent to multiplication in the $s$-domain, the Laplace transform of the $n^{th}$-order filter, $A_n(t)$, is given by:

$$L[A_n(t)] = \left[ \frac{b}{s + b - j\omega_0} \right]^n. \quad (5.18)$$

The Laplace transform of the derivative filter follows from the derivative theorem (Bracewell, 1986): $L[dA_n(t)/dt] = sL[A_n(t)]$. Here, the third-order filter and its derivative are implemented. The $z$-transforms of the band-pass filter and its derivative can both be expressed in the following form:

$$H_3(z) = q^3 \left( \frac{m_0 + m_1 z^{-1} + m_2 z^{-2} + m_3 z^{-3}}{1 + 3rz^{-1} + 3rz^{-2} + r^3z^{-3}} \right), \quad (5.19)$$

where $q = b/(b-j\omega_0+2)$, $r = (b-j\omega_0-2)/(b-j\omega_0+2)$, and the values of the $m_i$ are
given in table 5.1.

<table>
<thead>
<tr>
<th></th>
<th>( m_0 )</th>
<th>( m_1 )</th>
<th>( m_2 )</th>
<th>( m_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>band-pass</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>derivative</td>
<td>2</td>
<td>2</td>
<td>-2</td>
<td>-2</td>
</tr>
</tbody>
</table>

**Table 5.1:** Parameters for the discrete-time transfer function of the third-order filter and its derivative according to equation 5.19.

It should be noted that the transfer functions of the band-pass filter and its derivative share the same denominator, allowing a further reduction in computational storage. Discrete implementations follow from the \( z \)-transform (figure 5.5). We realise the common feedback path as the cascade of three first-order sections to ensure stability and to reduce coefficient sensitivity compared to the direct form (Tam et al, 1987; Nayeri & Jenkins, 1989).

**Temporal adaptation**

The goal of adaptation as discussed here is to minimise the error in image velocity measured from a small number of spatiotemporal filters. Assuming the least-squares solution to image velocity given in equation 5.10, we define the state of a set of filter kernels, \( v_f \), by their local velocity selectivity in a similar manner:

\[
\mathbf{v}_f = (\mathbf{K}^T \mathbf{W} \mathbf{K})^{-1} \mathbf{K}^T \mathbf{W} \mathbf{\Omega} = \mathbf{C} \mathbf{\Omega},
\]  

(5.20)
Recursive implementation of the third-order filter and its derivative expressed by the transfer function in equation 5.19 (see also table 5.1). The denominator of the transfer function (the feedback path of the filter) is implemented as the cascade of three first-order filters, each with transfer function \((1+rz^{-1})^3\). The spatially filtered image signal, \(I(t)\), passes through the feedback path of the filter, emerging as \(D(t)\). \(D(t)\) is then passed through the feedforward path to produce the overall response of the filter, \(R(t)\).
where $K$, $\Omega$ refer to the spatial and temporal frequency tunings of the filter kernels. In this formulation, the matrix $C$ is a constant because only the temporal parameters are adapted.

The adaptive scheme is posed in terms of a single-flow model of local image translation. In frequency space, all the power in a translating image lies on a plane through the origin (figure 5.6). Here, filters adapt their temporal centre-frequency tunings towards the plane of power in frequency space corresponding to the measured velocity of the local image signal, thereby tending to maximise the signal-to-noise ratio of their outputs (see figure 5.1). The centre frequency tuning of the $n^{th}$-order cascaded kernel, $A_n(t)$ (see equations 5.15, 5.16), is simply the temporal modulation frequency, $\omega_0$. Thus, to lie on the plane of power of the image signal, the centre-frequency temporal tuning of any filter tuned to a spatial frequency of $k_0$ must satisfy:

$$\omega_0 = v \cdot k_0.$$ (5.21)

Adjusting the parameter $\omega_0$, and hence the weights $q$ and $r$ in the IIR implementation (see equation 5.19), controls the peak temporal frequency tuning of the filter. Here, the value of $\omega_0$ is under adaptive control in an effort to minimise the distance of the peak tuning of the filter from the plane of power of the local image signal.
Figure 5.6: Spatially one-dimensional schematic diagram of a set of filters adapting their temporal tunings to lie on the line (in two spatial dimensions: plane) in frequency space corresponding to the velocity of the local image signal.
Dynamics of Adaptation

At each time step, measured velocity is fed back to adapt the filters. The feedback mechanism implemented is a vector formulation of the LMS algorithm of Widrow and Stearns (1985). Treating the locally measured velocity as the desired response vector, the temporal state of the filter kernels is adapted on the basis of the difference equation:

\[ e = v - v_f = v - C\Omega \]  \hspace{1cm} (5.22)

where \( e \) is a vector error signal. Following Widrow and Stearns (1985), we take the instantaneous value of the squared error, \( |e|^2 \), and minimise it with respect to \( \Omega \) to arrive at a gradient descent update equation:

\[ \Omega_{t+1}^T = \Omega_t^T - \eta_t \nabla_i \Delta \cdot |e|^2, \]  \hspace{1cm} (5.23)

\[ \Omega_{t+1}^T = \Omega_t^T - \eta_t e_t^T C, \]  \hspace{1cm} (5.24)

where \( \eta \) controls the rate of adaptation. Without noise, when \( \eta = 1.0 \), the adaptive
scheme converges in one iteration. In the simulations described below, using noisy input images, \( \eta \) is set at 0.1 to ensure convergence.

**Results**

The performance of the adaptive scheme described above is illustrated with results from two image sequences, one synthetic and one real. The synthetic diverging tree sequence (Barron et al., 1994; Fleet & Langley, 1995) has a known 2-D motion field, allowing a quantitative comparison of the performance of the adaptive and non-adaptive versions of the algorithm at various levels of additive white spatiotemporal noise. Results are presented graphically in the form of needle diagrams and confidence maps, and numerically as error statistics for the synthetic images.

**Implementation**

The spatial filters implemented are complex band-pass Gabor filters and their derivatives. The filters are tuned to each of 6 spatial orientations, with centre-frequency spatial tunings of 0.2 cycles per pixel and envelope standard deviations of 2.5 pixels. Each spatial filter feeds into a fixed low-pass temporal filter, as described in equation 5.15 but with \( \omega_0 = 0 \), and a pair of band-pass temporal filters of equal and opposite peak tuning. Under the adaptive scheme, the centre-frequency tunings of the band-pass temporal filters are adapted from an initial value of zero. In the non-adaptive scheme, used for comparison, the
temporal tunings of the band-pass filters are set to 0.2 cycles per frame, equivalent to a velocity tuning of 1 pixel per frame in the preferred direction. The temporal filter time constant (see equation 5.15), \( b^1 \), is set at 1.25 frames, giving an implicit delay of 3 frames in measurements of velocity. Following Fleet and Langley (1995), the spatiotemporal window, \( W(x,t) \), in equation 5.9 is a Gaussian in space (with a standard deviation of 1.2 pixels) and an exponential in time (with a time constant of 3.33 frames). Under the adaptive scheme, \( \eta \) (see equation 5.24) is set at 0.1.

**Error measures**

Following Fleet and Jepson (1990), velocity is viewed here as spatiotemporal orientation and error is measured as an angle in space-time. If velocity, \( v = (u,v)^T \), is represented as a unit vector, \( \mathbf{v} \), in space-time:

\[
\mathbf{v} = \frac{1}{\sqrt{u^2 + v^2 + 1}} (u,v,1)^T ,
\]

then the error, \( \psi_E \), between the correct velocity, \( \mathbf{v}_c \), and an estimate, \( \mathbf{v}_e \), is given by:

\[
\psi_E = \cos^{-1}(\mathbf{v}_c \cdot \mathbf{v}_e) .
\]
For results on synthetic image sequences a mean error and a density are quoted. The error is calculated as the mean angular error of all points satisfying an arbitrary confidence threshold. The density gives the percentage of points satisfying that threshold. The confidence measure used is the smallest eigenvalue of the spatial covariance matrix (see equation 5.10), which depends on the magnitudes of the spatial gradients and the range of their orientations.

**Measured flow fields**

During the diverging tree sequence the camera moves along its line of sight, the focus of expansion being the centre of the image. Speeds are up to 1.4 pixels per frame on the left of the image and 2.0 on the right. Figure 5.7a shows a frame from the sequence, and figure 5.8a the true flow field. The measured motion field under the adaptive scheme is shown in figure 5.8b to correspond closely to the true flow field. Encouragingly, the most noticeable deviation from the veridical field occurs in a region where confidence estimates are low. Confidence estimates are shown in figure 5.7b, with regions of high confidence being represented by bright points in the image.

To generate image noise, the noiseless sequence, \( I_0(x, t) \), was linearly scaled and added to the random variable, \( n(x, t) \), drawn from a distribution uniform over the range of values of the noiseless sequence. The noisy image, \( I_a(x, t) \), was created from the equation:
Figure 5.7: (a) Frame 20 of the diverging tree sequence; (b) confidence map obtained under the adaptive scheme with no added noise. The diverging tree sequence is a synthetic image sequence during which the camera moves along its line of sight towards the centre of the image. The scene consists of a highly textured tree on a less well-structured background. Regions of high confidence in velocity measurements, represented by bright points in the confidence map, are seen to correspond to areas of spatial structure in the image.
Figure 5.8: (a) Correct motion field for frame 20 of the diverging tree sequence; (b) measured motion field under the adaptive scheme with no added noise. The most noticeable errors occur in a region where confidence in the measured velocity is low (see figure 5.7b).
\[ I_n(x,t) = (1 - \alpha).I_0(x,t) + \alpha n(x,t) \]  

(5.27)

where 100\(\alpha\) is the percentage noise level. Numerical results from the fixed and adaptive schemes show that, at noise levels exceeding 5%, the adaptive scheme yields lower mean errors at a higher density of measurement (Table 5.2).

<table>
<thead>
<tr>
<th>Noise (%)</th>
<th>Fixed filters</th>
<th>Adaptive filters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>error (°)</td>
<td>density (%)</td>
</tr>
<tr>
<td>0</td>
<td>2.04</td>
<td>43.6</td>
</tr>
<tr>
<td>5</td>
<td>2.33</td>
<td>41.1</td>
</tr>
<tr>
<td>10</td>
<td>3.19</td>
<td>39.1</td>
</tr>
<tr>
<td>15</td>
<td>4.96</td>
<td>37.9</td>
</tr>
<tr>
<td>20</td>
<td>7.71</td>
<td>37.4</td>
</tr>
<tr>
<td>25</td>
<td>11.31</td>
<td>37.6</td>
</tr>
</tbody>
</table>

Table 5.2: Mean errors obtained from frame 37 of the diverging tree sequence, at various levels of added noise, using fixed and adaptive filters.

The performance of the adaptive scheme approaches that of its non-adaptive counterpart at low-levels of noise, whilst showing an advantage which increases with noise level for 10% noise and above. We would expect to be able to improve the performance of the fixed scheme by incorporating additional temporal filters, but even under the proposed IIR implementation this would involve a considerable increase in computational storage.
The rotating Rubik cube sequence (Barron et al, 1994) is a real image sequence in which a Rubik's cube is rotating anti-clockwise on a turntable. Figure 5.9a shows a frame from the sequence and figure 5.9b the flow field measured under the adaptive scheme after confidence thresholding.

Conclusion

An algorithm has been proposed for computing image motion from the phase output of centre-frequency adaptive IIR filters. The algorithm may find application in biomedical imaging, for example in perfusion or cardiac studies where image noise levels are potentially high. The main limitation of the technique is that some error is necessarily introduced into the velocity estimates by the process of adaptation itself, although for noisy image sequences this factor is outweighed by the improved signal-to-noise response of the adaptive filters. Indeed, as hardware developments allow faster frame-rates for real-time implementation of motion analysis algorithms, so the proportion of time during which significant adaptation occurs will decrease, and the benefit of superior image signal representation will outweigh the cost of adaptation even at low levels of noise.
Figure 5.9: (a) Frame from the rotating Rubik cube sequence; (b) measured flow field, after thresholding, under the adaptive scheme. The rotating Rubik cube sequence is a real image sequence in which a Rubik cube is rotating anti-clockwise on a turntable. Velocity measurements which satisfy the confidence threshold come from the spatially structured regions of the image, and clearly show the rotation present in the image sequence.
Chapter 6

Conclusions
This thesis investigated the problem of machine and biological motion perception from an adaptive systems perspective. Adaptive temporal filters were incorporated into two established image motion analysis algorithms, one each from machine and biological vision. The correlation-based Reichardt detector (Reichardt, 1961) was equipped with adaptive filters to account for existing electrophysiological data on motion adaptation in the insect lobula plate. The wider applicability of the model was tested by recording electrophysiologically from cells in the mammalian nucleus of the optic tract (NOT), and by investigating rapid adaptation to motion in human psychophysical observers. Adaptive temporal filters were also used to implement a phase-based scheme for image velocity measurement (Fleet & Jepson, 1989). The use of adaptive filters reduced the computational load of the phase-based scheme while maintaining performance on a synthetic test sequence. The adaptive scheme showed an advantage over its non-adaptive counterpart at high levels of noise as adaptation served to maximise the signal power in the output of the filters.

This concluding chapter discusses the main experimental findings and computational simulations presented in the thesis, and makes suggestions for further work stemming from that presented here.
Rapid neural adaptation to motion

Motion adaptation in insects

Insect visual systems have long been of interest to vision researchers because of the accessibility they afford to anatomical and physiological experimentation. The Reichardt model of motion detection is, at 35 years of age, one of the longest standing computational models in the neuroscientific literature, and a wealth of empirical work has gone into demonstrating its validity (see Reichardt, 1987; Egelhaaf et al, 1989). However, while its overall ability to predict the behaviour of motion-sensitive cells in the insect vision system is impressive, neither the original Reichardt model nor its subsequent elaborations have been able to account for the phenomenon of adaptation to motion.

Rapid neural adaptation to motion in insects is well documented (Zaagman et al, 1983; Maddess & Laughlin, 1985; de Ruyter van Steveninck et al, 1986; Borst & Egelhaaf, 1987; Maddess et al, 1991). The response of wide-field, motion sensitive neurons in the insect lobula plate to maintained motion at a constant velocity has consistently been found to rise rapidly to a peak at motion onset and then to decay over the course of a few seconds to a steady-state level. By stimulating only small areas of wide-field cells' receptive fields, it has been shown that the locus of adaptation is much smaller than the receptive field of the cell in which adaptation is measured. This has led to the conclusion that adaptation must be occurring in cells with smaller receptive fields, afferent to the wide-field cells
In chapter 2 of this thesis, an adaptive Reichardt detector model was proposed to account for the adaptation to motion of the H1 and HS cells of the insect lobula plate. Working within the existing correlation-based framework ensured that, at the starting point of the modelling, a great part of the phenomenology of insect motion vision could already be accounted for. Adaptation was implemented by feeding back the responses of motion-sensitive units to adjust the time constants, and hence the temporal frequency tunings, of temporal filters in the elementary motion detectors. In this way the adaptive model was able to account for the form of the lobula plate cells' responses to maintained motion, and for the progressive enhancement of sensitivity to changes in motion observed during adaptation (Maddess & Laughlin, 1985). In chapter 3, the adaptive model was modified to employ more biologically realistic components (Egelhaaf et al, 1989), and as a result was able to account for the directionality (Borst & Egelhaaf, 1987) and local nature (Maddess & Laughlin, 1985; de Ruyter van Steveninck et al, 1986) of motion adaptation. Thus, the widely accepted Reichardt detector model of insect motion vision was extended in this thesis to account for the phenomenon of adaptation to motion in lobula plate neurons.

Comparing motion adaptation in insects and mammals

Previous work has established parallels between the responses of insect lobula plate neurons and cells in the mammalian nucleus of the optic tract (NOT) (Ibbotson et
Lobula plate neurons are believed to be involved in the stabilisation of flight in insects (Shi Jian & Horridge, 1991), while the NOT in mammals is part of the accessory optic system involved in the control of eye movements (Simpson, 1984). Ibbotson et al (1994) showed that the responses to motion of NOT neurons in the wallaby, *Macropus eugenii*, were well accounted for by the Reichardt detector model. However, they also found marked adaptation to motion (Ibbotson & Mark, 1996) which could not be explained by existing formulations of the Reichardt framework. In chapter 3 of this thesis, electrophysiological studies carried out in collaboration with Ibbotson and Mark were presented which showed clear parallels between rapid neural adaptation in the wallaby NOT and that observed in the insect lobula plate. In both cases, it was shown that adaptation served to sacrifice absolute sensitivity to motion for differential sensitivity, i.e., sensitivity to changes in motion. The similarities in the behaviour of insect lobula plate and mammalian NOT neurons enabled the responses of both to be modelled using very similar versions of the adaptive Reichardt detector.

*Rapid adaptation to motion in humans*

Chapter 4 of the thesis considered the extrapolation of the electrophysiological findings to human vision. It was predicted that, if a human observer were making psychophysical judgements on the basis of the response of a population of cells with adaptive characteristics similar to those of insect lobula plate or wallaby NOT neurons, certain effects might be apparent. Firstly, the perceived speed of an adapting stimulus would decrease over the course of a number of seconds, in
analogy with the response decay observed physiologically. Secondly, the observer's sensitivity to small changes in the speed of an adapting stimulus would improve during the course of adaptation.

The two hypotheses were investigated psychophysically. To measure perceived speed, observers were required to discriminate between the speeds of an adapting stimulus and a short burst of a test stimulus at various adaptation durations. For each adaptation duration, a psychometric function was fitted to each subject's data, enabling a point of subjective equality (p.s.e.) to be determined. The p.s.e. was taken to be the perceived speed of the stimulus for that observer. To measure sensitivity to velocity change, observers were required to report the presence or absence of a brief modulation of velocity over a range of adaptation durations. The proportion of correct responses at each adaptation duration was taken as a measure of the differential motion sensitivity of the observer.

The psychophysical results supported the hypotheses. For each of three observers, perceived speed was found to decrease as a function of adaptation duration, while differential motion sensitivity improved. Thus, in analogy with the findings from insect and mammalian physiology, it was concluded that rapid adaptation to motion occurs in the human visual system, and that it serves to enhance differential sensitivity to motion at the cost of an accurate representation of absolute speeds.
Adaptive temporal filters for the phase-based measurement of image velocity

Chapter 5 described the use of adaptive temporal filters in a phase-based algorithm (Fleet & Jepson, 1989) for the efficient measurement of image velocity. The filters' centre-frequency temporal tunings were adaptively modified on the basis of measured velocity to optimise the representation of image motion. In computer simulations, the adaptive scheme was shown to provide accurate estimates of image velocity in both a synthetic and a real image sequence.

The phase-based scheme of Fleet & Jepson (1989) was taken as the starting point for the application of adaptive temporal filtering to machine vision as, in a comparative study, phase-based methods had been found to produce the most reliable measurements of image velocity (Barron et al, 1994). Adaptation was introduced into the phase-based scheme in order to improve efficiency by reducing the number of temporal filters required, and to increase accuracy by maximising filtered image signal power.

The performance of the scheme was illustrated with results from two image sequences, its degradation in the presence of additive noise investigated, and comparisons made with a non-adaptive version of the phase-based scheme. Standard image sequences were used to allow comparison with future results from elsewhere. Here, the accuracy of the adaptive scheme was shown to approach that of its non-adaptive counterpart on a noiseless synthetic image sequence, while demonstrating a considerable advantage at moderate and high levels of noise.
It is hoped that the algorithm proposed here may find application in biomedical imaging, for example in profusion or cardiac studies where noise levels are potentially high. The main limitation of the technique is that some error is necessarily introduced into the velocity measurements by the process of adaptation itself, although for noisy image sequences this factor should be far outweighed by the increased signal power in the output of the adaptive filters. Indeed, as hardware developments allow faster frame-rates for real-time implementation of motion analysis algorithms, so the proportion of time during which significant adaptation occurs will decrease, and the benefit of superior image signal representation should outweigh the cost of adaptation even at very low levels of noise. It is speculated that, in time, the use of adaptive filters for motion analysis will become commonplace in machine vision applications.

The role of adaptive motion analysis in machine and biological vision

Both the machine vision and biological modelling sections of this thesis employed adaptive temporal filters in a functional role. However, the nature of adaptation in the two cases was rather different. In the phase-based machine vision scheme, the filters adapted their centre-frequency temporal tunings to match the spectral composition of the image sequence. In the biological vision systems studied, the response to maintained motion was found to decay over time as the temporal filters adapted their peak centre-frequency temporal tunings to higher levels, away from the adapting temporal frequency.
It is instructive to compare the adapted steady-state response to motion under the
two schemes in terms of the range of possible responses of the artificial filters and
real neurons. Adaptation under the machine vision scheme served to maximise the
response to the prevailing image motion, while in the biological systems the
adapted steady-state response tended to be at a level near the middle of the
response range. This desaturation of the biological cell response ensured that, in
the adapted steady-state, the response to the adapting stimulus was at a point on
the upward slope of the temporal frequency response function where its gradient
was large. The change in response to a change in motion would thus also be large,
and the differential motion sensitivity of the cell would be enhanced from the
unadapted state.

What is at the root of the difference between the two schemes? The machine
vision scheme accurately measures image velocity, and adapts to optimise the
detection of motion. Adaptation in the biological systems, on the other hand,
sacrifices absolute sensitivity to motion for enhanced differential sensitivity. The
difference lies in the constraints on the representation of motion information in the
two cases. Cells in biological systems are only able to represent information with a
small number of distinguishable states, and are prone to failing to spike or to
spiking inappropriately. In signal processing parlance, motion information is
transmitted down neural channels of limited bandwidth which are susceptible to
transmission errors. Adaptation serves to make optimum use of these limited
resources, increasing the effective dynamic range of neural channels. In the
machine vision system the transmission of information is essentially perfect, since
all that is involved is reading from and writing to memory in a digital computer.

Thus, in the machine vision case, it is the detection rather than the transmission of motion information which is the limiting factor in determining accuracy. While accurate image velocity measurement is an attainable goal in machine vision, the limited number of response states available to a real neuron for coding information make an accurate representation of absolute velocity values unrealisable biologically. However, by shifting its operating range appropriately, a motion-sensitive neuron is able to maintain sensitivity to changes in motion over a wide range of baseline speeds in an analogous fashion to the way photoreceptors are able to respond to variations in light intensity around an ambient level. In their respective ways, both machine and biological vision systems can be seen as trying to perform optimally the joint task of detecting and transmitting motion information. In the machine vision case, the "transmission" problem is a trivial one, and adaptation can thus be used to optimise detection. In the biological systems it is the transmission of motion information which is the limiting factor, and hence optimal allocation of neural response states is of paramount importance.

Further questions

During the course of the work towards this thesis, questions have arisen which it has not yet been possible to answer. In the realm of electrophysiology, there are predictions of the adaptive Reichardt model to be tested, and parallels between insect and mammalian motion adaptation which merit further investigation. There are questions about human adaptation to motion and other stimulus dimensions to
be investigated psychophysically, and possible extensions of the adaptive machine vision scheme to deal with cases of motion transparency.

*Electrophysiology*

The adaptive behaviour of insect lobula plate neurons and mammalian nucleus of the optic tract (NOT) cells is in many regards very similar. However, a qualitative difference in the directionality of adaptation was described in chapter 3 of this thesis. While both insect lobula plate and mammalian NOT neurons adapt strongly to motion in the preferred direction, a difference was found in the effect on time constant adaptation of temporal modulation not associated with preferred motion. In a previous study, Borst and Egelhaaf (1987) reported significant adaptation of the time constant of the impulse response decay of fly H1 cells to motion in the anti-preferred direction and to counter-phase flicker. However, cells in the wallaby NOT were found in chapter 3 only to adapt strongly to motion in the preferred direction (figure 3.6). Before concrete conclusions about the significance of such a difference can be drawn, it would be wise to replicate both studies at a range of temporal frequencies. It would be particularly interesting to use a range of temporal frequencies extending beyond the cells' peak tunings, as the adaptive Reichardt detector model predicts that it should be possible to observe in insects the U-shaped function of adapted time constant against adapting temporal frequency found for wallaby NOT cells (Figure 3.4).
Chapter 4 of the thesis investigated rapid adaptation to motion in human vision. It was found that the perceived speed of an adapting stimulus decreased as a function of adaptation duration (Figure 4.2), while differential velocity sensitivity improved (Figure 4.6). However, what is not clear from the experiments reported here is whether it is stimulus motion per se which is driving adaptation, rather than the temporal modulation of intensity associated with that motion. One way to distinguish between the two possibilities would be to replicate the experiments using counter-phase flickering gratings rather than translating ones. If an effect of adaptation on perceived flicker rate were found, then it could be concluded that adaptation was governed, at least to some degree, by temporal modulation. If there were no effect of adaptation on perceived flicker rate, then it could be concluded that it was motion per se that was driving adaptation.

It would also be of interest to conduct, in the contrast domain, experiments analogous to those described here. By modulating the contrast of static luminance-defined patterns, contrast-modulated stimuli can be created. The effect of adaptation on the detection of such contrast modulations is germane to the question of how contrast modulations are detected by the visual system (Langley et al, 1996), as well as being of interest of itself. Greenlee and Heitger (1988) found that adaptation to a high contrast luminance-defined pattern improved contrast discrimination around the adapting level, suggesting that the detection of contrast modulations of a pattern should be enhanced by prior adaptation to that pattern.
However, subsequent work by Maattanen and Koenderink (1991) has failed to show such an effect, finding instead that contrast discrimination thresholds remain unchanged following adaptation.

**Machine vision**

As it stands, the adaptive phase-based scheme for image velocity measurement forwarded in chapter 5 assumes a model of local image translation. However, such a single-flow model of image motion is inadequate to deal with motion at object boundaries and cases of transparent motion. In such situations, there can be multiple components of optical flow at each point in the image, violating the assumption of local image translation. Constraint equations for multiple motions have been derived by cascading the differential operator for a single-flow constraint (Shizawa & Mase, 1991a), and a phase-based realisation of this approach has been proposed (Langley et al, 1992). Geometrically, all the power in a translating image lies on a plane through the origin in frequency space, while all the power in an \( n \)-flow transparent situation lies on \( n \) such planes. In order that transparency is not predicted when only a single flow is present, models capable of extracting multiple flows require a method for determining the degree of multiplicity of the flow field. Shizawa and Mase (1991b), for example, propose a comparison of the margin energies of all possible multiplicities up to an arbitrary degree to establish the multiplicity of the model which gives the most stable fit to image data.

Under the scheme of chapter 5, the filters adapt their temporal tunings towards the
plane in frequency space corresponding to local measured image velocity under an assumption of local image translation. To extend the scheme to deal with multiple motions would require either that the set of filter kernels adapted to lie on multiple planes in frequency space, or that adaptation was to a single plane even in cases where a multiple-flow model was fitted. In the first case, there is the problem of which filters adapt to which flow; in the second, there is the question of which plane all the filters should adapt to. Dealing with these issues would enable the adaptive image velocity measurement scheme presented here to be generalised to situations where the assumption of local image translation breaks down.
References


visual system", *Biological Cybernetics*, **55**, 367-375.


Lakatos, I. (1970). "Falsification and the methodology of scientific research 
programmes", in I.Lakatos & A.Musgrave (Eds.), *Criticism and the growth of 
knowledge* (pp.91-196), Cambridge University Press.

& G.Currie (Eds.), *The methodology of scientific research programmes. 
Philosophical papers of Imre Lakatos* (Vol 1, pp.102-138), Cambridge University 
Press.


*Proceedings of the 9th Israeli Symposium on AI and Computer Vision*, 255-264.


Rogers, B.J. & Graham, M. (1982). "Similarities between motion parallax and


Strausfeld, N.J. (1989). "Beneath the compound eye: neuroanatomical analysis and physiological correlates in the study of insect vision", in Stavenga, D.G. & Hardie,
R.C. (eds.), *Facets of vision* (pp. 317-359), Springer-Verlag, Berlin.


