The resolution of visually guided behaviour

Citation for published version (APA):

Kortmann, L. J. (2003). *The resolution of visually guided behaviour*. [Doctoral Thesis, Maastricht University]. Datawyse / Universitaire Pers Maastricht. <https://doi.org/10.26481/dis.20030704lk>

Document status and date:

Published: 01/01/2003

DOI:

[10.26481/dis.20030704lk](https://doi.org/10.26481/dis.20030704lk)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

The resolution of visually guided behaviour

The resolution of visually guided behaviour

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit Maastricht,
op gezag van de Rector Magnificus,
Prof. dr. A.C. Nieuwenhuijzen Kruseman,
volgens het besluit van het College van Decanen,
in het openbaar te verdedigen
op vrijdag 4 juli 2003, om 12.00 uur

door

Laurens Jan Kortmann

Promotores: Prof. dr. H.J. van den Herik
Prof. dr. E.O. Postma

Leden van de beoordelingscommissie:

Prof. dr. J.M.J. Murre (voorzitter)
Prof. dr. M. Egelhaaf (Universität Bielefeld, Duitsland)
Prof. dr. ir. A. Hasman
Prof. dr. H. Kingma
Prof. dr. B.J. Wielinga (Universiteit van Amsterdam)

Dissertation Series No. 2003-09

ISBN 90-5278-382-9

Universiteit Maastricht

©2003 Rens Kortmann

Cover design: Ellen Oosterwijk

Cover photos: H.L. Leertouwer and National Eye Institute, National Institutes of Health.

Preface

Why do the eyes of insects look so different from those of humans? How do flies always manage to escape when you try to swat one? Do they really perceive time slower than we do? And how do they decide in what direction to flee? This sort of questions have fascinated me ever since my first biology classes.

Already as a child I marvelled at pictures of animals and their various appearances. During my undergraduate studies at university I was introduced into the biophysics of visual detection and the psychology of decision making. Then I learnt to use computer simulations as powerful modelling tools to understand the astonishing creations of nature. In my graduation project I trained a mobile robot to resemble the ways in which female crickets distinguish between the different calling songs of males, an ability that many humans may not attain. Inspired by the success of the project I decided to pursue a Ph.D. title. For this I wanted to investigate the control mechanisms of animal behaviour. My goal was to construct computer simulations to understand the intricacies of the manners in which animals use their eyes to cope with their complex environment.

I was most fortunate that my supervisors, Jaap van den Herik and Eric Postma, allowed me to pursue my goals and I owe many thanks to their support and commitment. They were always there to answer my questions and to discuss my work, also when political headwinds must have caused their thoughts to drift astray.

I would like to thank my colleagues at the department of computer science for their cooperation and help. In particular, I'd like to express my appreciation for the support offered by the secretarial staff, Marlies van der Mee, Joke Hellemons, Hazel den Hoed, and Martine Tiessen, who took care of many administrative tasks in my pursuit of a Ph.D. and who performed an excellent job in 'managing' my supervisors. Furthermore I thank Jo Beerens and Peter Geurtz for computer support; Jeroen Donkers for his help on Latex formatting; Erik van der Werf for many Linux issues; Ida Sprinkhuizen-Kuyper for co-authoring papers and for proofreading my work; Evgueni Smirnov for acting as a mental coach; and my fellow preachers of situatedness — Michel van Dartel, Paul Vogt, Joyca Lacroix, and Dirk Bollen — for inspiration, proofreading, tea, and lots of fun with our robots.

Outside the University of Maastricht I feel indebted to Henk Mastebroek and Erwin Mulder who helped me getting started with the biophysics literature and who co-organised our workshop on robots and vision in January 2002 at the University of

Groningen. Many thanks go to Eric Warrant, Roland Kern, and David O'Carroll for explaining the relevant facts of invertebrate vision. The help of Lieke Zwanenburg in correcting the English was appreciated greatly. Moreover, the exchange of ideas with the members of 'Emergentsia', a discussion group of Dutch Ph.D. students, is kindly acknowledged (Bram Bakker, Martijn Brinkers, Paul den Dulk, Hendrik Jan Hoeve, Edwin de Jong, Michiel de Jong, Dick de Ridder, David Tax).

Outwith the academia I was happy to unload the stress and strain of writing a Ph.D. thesis on the floors of Maastricht's drama theatres. Here my sincere thanks go to the boards and members of student theatre company 'Alles is drama', theatre company '101punt', melodic performance group 'Boddie', Bas Banning, Raoul Boers, and many other people from the scene.

I am grateful to my parents for their unconditional support and faith. I feel happy to know that I can always rely on them. Finally, I thank Ellen for loving me through it all.

Rens Kortmann
Maastricht, 4 July 2003

Contents

Preface	v
Contents	vii
1 Introduction	1
1.1 Spatial resolution and temporal resolution	2
1.2 A trade-off between spatial and temporal resolution	2
1.2.1 Insects and vertebrates	2
1.2.2 Fovea and periphery	6
1.2.3 Photography	6
1.3 Which are the causes of the trade-off?	6
1.4 Thesis outline	7
2 Studying the resolution of visually guided behaviour	9
2.1 A taxonomy of approaches	9
2.2 Four levels of detail	11
2.2.1 The phenomenological level	11
2.2.2 The systems level	12
2.2.3 The neural level	14
2.2.4 The cellular level	16
2.3 Two classes of research	17
2.3.1 Animal research	17
2.3.2 Modelling research	17
2.4 Two types of investigation	18
2.4.1 Information theory	18
2.4.2 Situated modelling	19
2.5 Situated systems modelling	20
3 Scrutinising the trade-off	23
3.1 Control of visually guided behaviour	24
3.1.1 Two principal control systems	24
3.1.2 Position-detection systems	24

3.1.3	Motion-detection systems	25
3.2	Candidate causes of the trade-off	26
3.2.1	Five determinants of the trade-off	27
3.2.2	Light sensitivity	27
3.2.3	Contrast sensitivity	30
3.2.4	Motion sensitivity	31
3.2.5	Heat production	32
3.2.6	Energy consumption	33
3.3	How to find generic causes	34
4	The resolution of position-detection systems	37
4.1	Aim of the chapter	37
4.1.1	Chapter survey	38
4.2	A model of target tracking	38
4.2.1	Types of target tracking	38
4.2.2	Components of the model	39
4.2.3	Implementation	40
4.2.4	Position-detection mechanism	42
4.2.5	Section summary	44
4.3	Optimisation with an evolutionary algorithm	44
4.3.1	The fitness function	45
4.4	Evolution of a trade-off by light sensitivity and motion sensitivity . .	46
4.4.1	Implementation in the model	46
4.4.2	Experimental details	47
4.4.3	Analytical predictions	47
4.4.4	Results	48
4.4.5	Discussion	50
4.4.6	Section conclusion	52
4.5	Evolution of a trade-off by heat production and motion sensitivity .	53
4.5.1	Implementation in the model	53
4.5.2	Experimental details	54
4.5.3	Analytical predictions	54
4.5.4	Results	55
4.5.5	Section conclusion	56
4.6	Evolution of a trade-off by energy consumption and motion sensitivity	56
4.6.1	Implementation in the model	56
4.6.2	Analytical predictions	57
4.6.3	Section conclusion	57
4.7	Comparison to biological systems	57
4.7.1	The role of the target's angular velocity	58
4.7.2	The role of the target's illumination	58
4.7.3	The role of heat dissipation	59
4.7.4	The role of energy supply	59
4.7.5	Section summary	60

4.8	Chapter discussion and conclusions	60
4.8.1	Application to biological systems	60
4.8.2	Limitations and possible improvements of the model	60
4.8.3	Chapter conclusions	61
5	The resolution of motion-detection systems	63
5.1	Aim of the chapter	63
5.1.1	Method	64
5.1.2	Chapter survey	64
5.2	A situated model of gaze stabilisation	65
5.2.1	Components of the model	65
5.2.2	The augmented Reichardt correlator	66
5.2.3	The frequency-dependent response function	68
5.2.4	The velocity-dependent response function	72
5.2.5	Section summary	74
5.3	How visual resolution sets motion sensitivity	75
5.3.1	The relation in the model	75
5.3.2	The role of spatial natural image statistics	76
5.3.3	The role of negative response	77
5.3.4	The role of pass-band width	79
5.3.5	Section conclusion	81
5.4	The optimal motion sensitivity	81
5.4.1	Signal-to-noise ratio optimisation	82
5.4.2	Gain-control optimisation	85
5.4.3	Combined SNR and gain-control optimisation	94
5.4.4	Section conclusion	96
5.5	How visual resolution sets light sensitivity and contrast sensitivity	98
5.5.1	Response amplitude as a function of resolution	99
5.5.2	Response amplitude as a function of illumination and contrast	101
5.5.3	Sensitivity as a function of resolution	104
5.5.4	Section summary	105
5.6	The optimal light sensitivity and contrast sensitivity	106
5.6.1	Optimisation of SNR and latency	107
5.6.2	The optimal resolution scale	108
5.6.3	Section summary	111
5.7	A trade-off in motion-detection systems	113
5.8	Chapter discussion and conclusions	114
5.8.1	Comparison to biological systems	115
5.8.2	Comparison to other models	116
5.8.3	Chapter conclusions	118

6	The importance of being situated	121
6.1	Comparison to controlled behavioural studies	121
6.1.1	Comparability of the results	122
6.1.2	Availability of relevant data	122
6.1.3	Provision of insight	122
6.2	Comparison to information theory	123
6.2.1	Maximising information capacity	123
6.2.2	Results	123
6.2.3	Conclusions	124
6.3	The importance of situated systems modelling	125
7	Generic causes of the trade-off	127
7.1	The determinants iterated	128
7.2	Additional determinants	128
7.3	Alternative occurrences of behaviour	129
7.3.1	Combined position-detection and motion-detection systems .	130
7.3.2	Multimodal control	130
7.3.3	Control of other types of visually guided behaviour	131
7.3.4	Consequences for the generalisation of the determinants . . .	132
7.4	Applying determinants to biological systems	132
7.4.1	A biological resolution data set	133
7.4.2	Predicting trade-off points	133
7.4.3	Matching position	135
7.4.4	Matching scale	136
7.4.5	Section summary	137
7.5	Applying determinants to artificial systems	138
7.6	The generic causes	138
8	Conclusions	141
8.1	Chapter conclusions summarised	141
8.2	Generic causes of the trade-off	143
8.3	Situated modelling: a general conclusion	143
8.4	Future research	144
Appendices		
A	Filter response functions	145
B	Stability of the update rule	147
References		151
Summary		167
Samenvatting		171

<i>Contents</i>	xi
Curriculum Vitae	175
SIKS dissertation series	177

Chapter 1

Introduction

“Every animal has ideas, since it has senses; it even combines those ideas in a certain degree; and it is only in degree that man differs, in this respect, from the brute.”

Jean-Jacques Rousseau (1712 – 1778), Swiss-born French philosopher.

Chapter contents: Visual resolution, a trade-off between spatial and temporal resolution, research question, thesis outline

Animals act in a visual world. Their visually guided behaviour is governed by their visual resolution. For instance, the eyes of human beings have a high resolution, which enables a behaviour that is more complex than that of *Nautilus* (a molluscan animal species). Such a mollusc has eyes with a much lower visual resolution (Fernald, 2000). Maximising the visual resolution is an important principle in the animal world. It is a significant way to attain a behaviour of high complexity. In this thesis we study the determinants of visual resolution. We aim at determining the generic causes of a trade-off between spatial and temporal resolution.

In this chapter we give some background information and formulate our research topic more precisely. In section 1.1 we distinguish two types of visual resolution: spatial resolution and temporal resolution. Section 1.2 provides some empirical evidence as observed by others that the spatial and temporal resolution of animals are never maximised simultaneously. This is a striking observation. A direct consequence might be that the visual resolution is subject to a trade-off. This idea leads us, in section 1.3, to the formulation of our research question. It is followed by an outline of the thesis in section 1.4.

1.1 Spatial resolution and temporal resolution

Visual resolution is generally characterised by two types: spatial resolution and temporal resolution. Below, we explain both types and show that maximising each of the resolution values is important to accomplish optimal visually guided behaviour.

Spatial resolution indicates the level of detail that a visual system can detect in a static scene. The importance of maximising the spatial resolution becomes clear from the following observation. Certain birds of prey employ high values of spatial resolution to detect and hunt their small prey from high altitudes. For instance, the African serpent eagle resolves spatial frequencies up to 120 cycles per degree (usually denoted as c/deg). In comparison, the human spatial resolution is 60 c/deg (Miller, 1979) which results in a less detailed sight. Obviously, maximising the spatial resolution is important to optimise visually guided behaviour.

Temporal resolution determines how well the visual system responds to changes in a scene. A high temporal resolution enables better detection of moving objects than a low temporal resolution does. The timely detection of moving objects is vital for many types of visually guided behaviour, such as the avoidance of approaching objects (Blanchard, Rind, and Verschure, 2001), or the pursuit of other animals (Collett, 1980). Therefore, maximising the temporal resolution is important as well to optimise visually guided behaviour.

1.2 A trade-off between spatial and temporal resolution

Remarkably, the simultaneous occurrence of high spatial resolution and high temporal resolution is never encountered in nature. Why do animals not adopt the best of both worlds? Studies of visual systems (*i.e.*, the eyes and the visual pathways in the brain) suggest that there exists a trade-off between spatial and temporal resolution (Kirschfeld, 1976; Land, 1981; Warrant, 1999). This means that high temporal resolution may only be attained at the cost of low spatial resolution and *vice versa*. Below we provide three examples of a trade-off. They are taken from a comparison between insects and vertebrates (subsection 1.2.1), a comparison between the fovea and the periphery of the retina (subsection 1.2.2), and from an artificial visual system, namely a photocamera (subsection 1.2.3). From the examples it will be derived that not much is known about the exact causes of the aforementioned trade-off.

1.2.1 Insects and vertebrates

Below we compare the spatial resolution of insects and vertebrates, followed by a comparison of the temporal resolution. Finally, we point at the existence of a trade-off.

Spatial resolution

It is well-known that the spatial resolution of some invertebrate eyes, *e.g.*, insects, is much lower than that of vertebrate eyes (Land, 1981). The differences between the spatial resolution of insects and that of vertebrates are caused by the differences in their eye design. Below we show the differences in eye design and then explain the differences in spatial resolution.

Insects have compound eyes whereas vertebrates possess single-lens or camera-type eyes. Both eye designs are depicted in figure 1.1 (a) and (b), respectively (adopted from Warrant (1999)). Since we focus on a comparison of the spatial reso-

(a) Insect eye

(b) Vertebrate eye

Figure 1.1: The eyes of insects and vertebrates. (a) a schematic drawing of a cross-section of a compound eye of insects and (b) of a single-lens eye of vertebrates. Both figures were adopted from (Warrant, 1999).

lution, rather than on an exhaustive description of the eyes, we constrain ourselves to providing a brief description only. Compound eyes (figure 1.1(a)) are composed of visual units called ommatidia (the collection of parts labelled as cl, cc, p, and rh). Each visual unit contains a facet lens (cl). In contrast, single-lens eyes (figure 1.1(b)) possess just one lens (l) for all visual units (ph).

The spatial resolution of eyes may be determined by the density of visual units (*i.e.*, anatomical resolution; Snyder, 1979). In compound eyes, a high density is obtained by employing long and narrow ommatidia: small facet lenses in large compound eyes lead to a high spatial resolution. In single-lens eyes, a high density results from closely packing the visual units.

Snyder (1979) pointed out that the size of lenses in an eye has a lower bound set by a so-called diffraction limit. Below this limit, lenses scatter the incoming light rays which leads to a blurred image formation. Due to diffraction limitation and

the small size of insect eyes, Land (1981) concluded that the spatial resolution of the compound eyes of insects (density: 1 visual unit per degree; Wehner, 1981) is much lower than that of vertebrate single-lens eyes (density: ± 20 visual units per degree; derived from Wandell, 1995). To illustrate the differences, Kirschfeld (1976) calculated the diameter of compound eyes needed to obtain the spatial resolution of the vertebrate single-lens eyes. He arrived at a diameter of 1 m. Figure 1.2 is a well-known illustration of this result.

In summary, due to the diffraction limit of their compound eyes, the spatial resolution of insects is much smaller than that of vertebrates.

Temporal resolution

Temporal resolution deals with changes in a visual scene, such as those caused by moving objects or by self motion. Laughlin (1981) stated on temporal resolution that certain insects developed a visual system of which the “temporal resolving power is, by comparison with [temporal resolution in] vertebrates [...], excellent” (pp. 222 – 223).

For instance, the flicker fusion frequency¹ of certain flies is 200 Hz (Tatler, O’Carroll, and Laughlin, 2000) and exceeds that of humans which is only 2 – 6 Hz (when dark-adapted; Toveé, 1996). The low flicker fusion frequency of humans is clearly illustrated by the fact that the human visual systems sees a motion picture projected at a rate of 25 frames per second as a smoothly moving image. This measure of visual resolution is termed behavioural resolution since it is derived from the behavioural response of organisms (*e.g.*, humans) to a visual test stimulus (*cf.* anatomical resolution above).

Further data of temporal resolution are provided in chapter 7. As a starting point for our research we follow Land (1981) who claims that the temporal resolution of many insects is higher than that of vertebrates.

A resolution trade-off

From previous investigations (Land, 1981) and our comparison it follows that, in general, the eyes of neither insects, nor vertebrates possess the maximal values of both the spatial resolution and the temporal resolution. In particular, we concur with the general opinion (Kirschfeld, 1976; Land, 1981; Warrant, 1999) that the data hint at the fact that the spatial and temporal resolution are traded against each other. The examples given show that the low spatial resolution in insects is caused by the design of their compound eyes. However, the differences between the temporal resolution of insects and that of vertebrates are not explained by this design. Moreover, the trade-off between spatial and temporal resolution in insects and vertebrates is our main topic. So far, the exact causes of this trade-off are unknown.

¹The frequency at which a flickering light source is detected as a continuous stimulus.

Figure 1.2: The size of compound eyes needed to obtain the spatial resolution of human single-lens eyes (from Kirschfeld, 1976). The size of the facet lenses has a lower bound due to diffraction limitations. Therefore, high resolution may only be obtained by increasing the eye size.

1.2.2 Fovea and periphery

As a second example we mention a trade-off between the spatial and temporal resolution of the fovea (central part of the retina) and of the peripheral parts of the retina.

In the human eye, many visual units in the fovea have a high spatial resolution, but a low temporal resolution (Livingstone and Hubel, 1988). The inverse is true for the visual units in the peripheral parts of the retina. A similar phenomenon is observed in the central and peripheral parts of the eyes of certain insects (Collett and Land, 1975).

These data suggest the existence of a trade-off between spatial and temporal resolution of the fovea and of the periphery. Barlow (1981) provided some reasons why the peripheral spatial resolution is lower than the foveal spatial resolution. Again, finding these reasons is not our aim. Here we focus on the exact causes of the trade-off between spatial *and* temporal resolution in the fovea and the periphery. These causes have not been identified so far.

1.2.3 Photography

As a third appearance of the trade-off between spatial and temporal resolution, we provide an example of an artificial visual system. In a photocamera the shutter time is a measure of temporal resolution, while the grain size of a photo-sensitive film sets the spatial resolution (Keller, 1993). To make photographs in dark environments, either the camera's shutter time or the grain size may be set to a high value. Moreover, to reduce motion blur while preserving light sensitivity, a decrease of the shutter time is accompanied by an increase of the film's grain size. In other words, the temporal resolution of the camera is traded for spatial resolution to obtain an optimal exposure of the film.

The causes of this trade-off are well known: light sensitivity and motion sensitivity. However, there may exist additional causes of a similar trade-off in other artificial visual systems, such as video cameras, motion sensitive chips, and other electronic devices. So far, no complete survey of causes has been established.

1.3 Which are the causes of the trade-off?

In section 1.2 we clearly expressed that at present the exact causes of the trade-off between spatial and temporal resolution are unknown. However, it is generally assumed that the trading takes place to optimise the visually guided behaviour of animals as a function of their life style (Warrant, 1999). For instance, it was found that animals which detect much image motion in their natural behaviour adapted their visual system to a high temporal resolution (O'Carroll *et al.*, 1996). Therefore, it is suggested that the trade-off is a result of adaptations to the visual ecology of animals, *i.e.*, to the visually guided behaviour of animals in their natural environment.

It is our aim to explain the existence of a resolution trade-off, *i.e.*, to determine its causes. From the suggestion mentioned above we arrived at the following research question:

“Which are the generic causes of a trade-off between the spatial and temporal resolution in visually guided behaviour?”

We define generic causes by three conditions that must be fulfilled. Generic causes (1) explain the existence of a trade-off in many occurrences of visually guided behaviour, (2) cause a trade-off in many different biological visual systems, and (3) cause a trade-off in artificial visual systems.

1.4 Thesis outline

The thesis is organised as follows. In chapter 2 we establish that situated modelling is a suitable approach to studying the resolution of visually guided behaviour. Our research method is presented and motivated in chapter 3. There we determine how situated modelling is employed to answer our research question. Then we study the trade-off between spatial and temporal resolution in two models of visually guided behaviour in order to find the generic causes. Chapter 4 describes the construction of the first model that is based on a position-detection system. The chapter also reports on our experiments and results with this model. Chapter 5 does the same for the second model, which is based on a motion-detection system. To indicate the importance of situated modelling, we compare our modelling results to those obtained with other research approaches in chapter 6. Chapter 7 determines the generic causes of the trade-off. Finally, in chapter 8 we provide the conclusions of our research and describe future research directions.

Chapter 2

Studying the resolution of visually guided behaviour

“I approach these questions unwillingly, as it wounds, but no cure can be effected without touching upon and handling them.”

Titus Livius (59 B.C. – A.D. 17), Roman historian.

Chapter contents: A taxonomy of approaches, levels of detail, classes of research, types of investigation, situated systems modelling

Studying the resolution of visually guided behaviour starts with deciding upon a research approach. This chapter works towards a suitable approach that may lead to a satisfactory answer to our research question.

In section 2.1 we propose a taxonomy, inspired by Egelhaaf and Borst (1993a), that classifies the current research approaches to studying visually guided behaviour along three dimensions. We discuss the dimensions in sections 2.2, 2.3, and 2.4. Based on the discussion we choose our research approach in section 2.5.

2.1 A taxonomy of approaches

The taxonomy of current research approaches is illustrated in figure 2.1. The three dimensions along which the current approaches to studying visually guided behaviour may be classified are: the level of detail, the class of research, and the type of investigation. We briefly discuss each dimension.

Figure 2.1: A taxonomy of research approaches (cf. Arbib, 1987; Shepherd and Koch, 1990; Egelhaaf and Borst, 1993a). Three dimensions are indicated along which approaches are classified: (1) the level of detail (from low to high: four levels); (2) the class of research (animal research or modelling research: two classes); and (3) the type of investigation (*e.g.*, information theory or situated modelling, dependent on the level and the class; in the figure they are only indicated for systems modelling to increase readability).

First, we distinguish four levels of detail to analyse and explain the visually guided behaviour of flies (see also Arbib, 1987; Shepherd and Koch, 1990; Egelhaaf and Borst, 1993a). We term them the phenomenological level (highest), the systems level, the neural level, and the cellular level (lowest). A high level yields global results (little detail) whereas a low level leads to rather specific results (much detail).

Second, at each level we distinguish pairs of classes of research: one class contains research approaches to studying animals; the other class contains research approaches to studying models of animals (see also Egelhaaf and Borst, 1993a). In figure 2.1 we show four pairs: (1) uncontrolled behavioural studies and phenomenological modelling; (2) controlled behavioural studies and systems modelling; (3) neural studies and neural modelling; and (4) cellular studies and cellular modelling.

Third, every class contains various types of investigation. In figure 2.1 we indicate two types for the class of systems modelling: information theory and situated modelling.

In summary, we propose a taxonomy that classifies the current approaches to studying the resolution of visually guided behaviour. The taxonomy is based upon three dimensions: the level of detail, the class of research, and the type of investigation.

2.2 Four levels of detail

Following Egelhaaf and Borst (1993a), we discuss four levels of detail: the phenomenological level (subsection 2.2.1); the systems level (subsection 2.2.2); the neural level (subsection 2.2.3); and the cellular level (subsection 2.2.4). For each level, we present examples of animal research and modelling research and identify the advantages and disadvantages of the level.

2.2.1 The phenomenological level

The phenomenological level of detail studies the observable behaviour of animals in the natural world. It aims to provide stimulus-response analyses of the behaviour, rather than a description of the mechanisms that control the behaviour. Therefore, the level views animals as ‘black-box systems’.

At this level, the animal research is represented by uncontrolled behavioural studies. The modelling research comprises phenomenological modelling.

Uncontrolled behavioural studies

The ideal approach to investigating visually guided behaviour in animals is by studying the animals that move freely in uncontrolled environments (cf. Tinbergen, 1951). For instance, Land and Collett (1974) used high-speed cameras to measure the behavioural patterns of freely flying flies. This approach is used to identify different occurrences of visually guided behaviour, such as target tracking and gaze stabilisation.

Phenomenological modelling

Phenomenological modelling describes high-level behavioural phenomena, such as the different occurrences of visually guided behaviours (*e.g.*, tracking, gaze stabilisation, grasping). For instance, Collett (1980) modelled the behaviour of chasing flies recorded by cameras as a closed-loop control system. The model calculations yielded information of, *e.g.*, the gain of the control system. However, his phenomenological model neither revealed the details of photoreception nor the transfer functions of control systems found in the brain of the fly.

Advantages and disadvantages

The advantages of approaches of the phenomenological level are their ability to yield the least biased measurements, since the uncontrolled set-ups interfere little with the natural behaviour of animals (*i.e.*, the measurements are ‘ecologically valid’; Gibson, 1979). However, as an important disadvantage, the phenomenological level cannot be used to study the trade-off between spatial and temporal resolution. Such a study requires the determination of the visual resolution from the anatomy of the eyes (anatomical resolution), or from the correlations between patterns of visual

input and motor output of the animal (behavioural resolution). In contrast, since phenomenological models describe the motor behaviour of animals recorded by external observers, the models neither determine the anatomical resolution, nor the behavioural resolution (subsection 1.2.1).

Recently, new methods were proposed that may be used to determine the visual resolution of animals from uncontrolled behavioural studies. These methods reconstruct the visual input of freely flying and walking flies (Schilstra and Van Hateren, 1998; Van Hateren and Snippe, 2001; Kern *et al.*, 2001a). After extending these methods to many other animals, uncontrolled studies may be used in the future to study the trade-off. At present we believe that, based on the disadvantages mentioned above, the approaches belonging to the phenomenological level are still not suitable to study the resolution trade-off adequately.

2.2.2 The systems level

The level of detail that we term the systems level (cf. Van Essen, Anderson, and Felleman, 1992) studies the exact input-output relations of biological systems that control visually guided behaviour. Research approaches at this level aim to analyse the processing of input patterns in identified brain areas to establish the frequency response function of the area (Oppenheim, Willsky, and Nawab, 1997). As is shown in figure 2.1, the systems level comprises controlled behavioural studies and systems modelling.

Controlled behavioural studies

In general, controlled behavioural studies are employed to derive visual resolution from the behaviour of animals (cf. Mastebroek, Zaagman, and Lenting, 1980; Anderson and Burr, 1985). Such studies determine the exact correlations between the visual input and either the behavioural or the perceptual output of animals.

Srinivasan and Bernard (1975) distinguished two main ways of employing controlled behavioural studies to measure visual resolution: (1) by determining the minimal angular separation at which two objects can be discriminated and (2) by using Fourier analysis to determine the amplitude response function of the visual system. Using the first way, Barlow (1952) predicted the theoretically minimal separation angle (*i.e.*, spatial resolution) of compound eyes. Moreover, this use of controlled behavioural studies yields the minimal size of objects to be detected. Using the second way, animals are placed in an artificial set-up to measure how controlled spatio-temporal input patterns influence their behaviour. Examples of such artificial set-ups are: a Y-maze globe (Reichardt, 1961), a flight simulator (Fermi and Reichardt, 1963), and a treadmill (Götz, 1975).

An illustrative example of the second way (*i.e.*, the use of Fourier analysis) is provided by Schmid and Wildsoet (1998) who placed chicks in front of a moving screen onto which moving striped patterns were shown. Normally, the eyes follow the moving pattern involuntarily. However, when the stripes were too narrowly

spaced (high spatial pattern frequency) the retinal image of the pattern was blurred due to optical and neural filtering. As a result, the pattern motion was not detected properly and no eye movements could be observed. Using this set-up, the maximal spatial frequency was determined for which the behaviour was shown. The frequency obtained in this method corresponds to the spatial resolution (Reinecke and Cogan, 1958).

Systems modelling

According to figure 2.1, the approach of systems modelling is paired with controlled behavioural studies. Both approaches yield quantitative input-output relations. Systems modelling provides the mathematical equations that describe the underlying control systems of visually guided behaviour in animals.

An important example of a systems model is the Reichardt correlator that emulates the response of biological motion detectors (Hassenstein and Reichardt, 1956; Reichardt, 1961). Figure 2.2.2 depicts a basic version of a Reichardt correlator model. The model comprises two input sensors (the ‘cups’, separated at distance

Figure 2.2: The basic Reichardt correlator model for biological motion detection.

$\Delta\phi$), a time delay Δt , and a multiplication unit which models a correlation process. The motion sensitivity of the system is explained as follows. The figure shows a target light source T moving at a velocity V_T . When T moves to the right as indicated, the left sensor is stimulated before the right sensor is. However, the left sensor's response is delayed by a time period Δt . If the fraction of the sensor distance $\Delta\phi$ and the delay Δt equal the velocity of the light source, *i.e.*, $V_T = \Delta\phi/\Delta t$, the responses arrive at the multiplicative unit simultaneously. Therefore, the correlator's output is maximal when the velocity is matched to the spatial and temporal properties of the correlator. However, if this is not the case, *e.g.*, if the velocity is opposite as indicated or if the velocity is higher or smaller than the quotient, the output is low. Therefore, the Reichardt correlator is directionally and selectively responsive

to image motion, which corresponds to the response of biological motion-sensitive systems (Reichardt, 1961).

Advantages and disadvantages

From the paragraph above it follows that, as an advantage, controlled behavioural studies yield precise input-output relationships of the system of interest. A disadvantage of such studies is that the experimental set-ups may distort the behaviour investigated. The artificial conditions in which the studies are performed are prone to introduce a bias in the behaviour (Zanker and Zeil, 2001a). We mention two possible occurrences of such a bias. First, by simplifying the experimental set-up, some experiments focus on open-loop control of behaviour, *i.e.*, the animal in the experiment does not experience the consequences of its actions. In chapter 5 we show that the optimal visual resolution for open-loop and closed-loop (*i.e.*, natural) behaviour are rather different (see also Götz, 1975; Heisenberg and Wolf, 1993). Therefore, the open-loop studies introduce an experimental bias. Second, the artificial set-ups of controlled behavioural studies (striped input patterns, monitor screens, controlled lighting conditions, etc.) may also affect the natural behaviour of animals and thus introduce an experimental bias (Gibson, 1979).

In conclusion, systems modelling is suitable to study the resolution trade-off. However, the approach may introduce a bias in the results due to the artificial experimental set-ups.

2.2.3 The neural level

The neural level of detail focuses on the operation of neural pathways in the brains of animals. Below we provide examples of neural studies of animals and the corresponding neural modelling approaches to arrive at the advantages and disadvantages.

Neural studies

To determine the spatial and temporal resolution of biological neural units (*e.g.*, the spatio-temporal receptive fields of ganglion cells), approaches such as staining, electrophysiology, and optical studies are employed (cf. Hausen, 1982; Shepherd, 1990; Felleman and Van Essen, 1991). We mention the following three examples. First, the staining of visual neurons by intercellular dye injections provides maps of neural morphology (Strausfeld, 1976). From the maps, the spatial arrangement of visual neurons (*e.g.*, photoreceptors and ganglion cells; Scholes, 1976) can be measured to yield the anatomical resolution. Second, intracellular electrical recordings produce spatio-temporal receptive fields of neurons, from which the spatial and temporal resolution can be derived (*e.g.*, Shapley, 2000). Third, the reflective layer in the compound eyes of certain insects is used to determine the distribution of photoreceptors optically (Qiu *et al.*, 2002) from which spatial resolution can be calculated.

Neural modelling

Neural modelling has been applied in many domains ranging from biological signal processing (see *e.g.*, Softky and Koch, 1993) to abstract neural computation, such as function approximation (McClelland and Rumelhart, 1991) or principal components analysis (PCA; Oja, 1995). To study the resolution of biological visually guided behaviour we found that the approach of biological neural network modelling (Arbib, 1995) is most suitable. In this approach, the spatial and temporal characteristics of biological neural networks are represented by mathematical equations to construct a computer model. For instance, Hennig and Funke (2001) modelled the vertebrate retina as a computer simulation. The model emulates the properties of biological retinas (including visual resolution). Therefore, biological neural network models may be used to understand the trade-off.

Advantages and disadvantages

The examples above illustrate the advantage of neural approaches: they yield the precise spatial and temporal resolution of individual neurons or groups of neural cells.

Since most neural approaches are controlled, they suffer from the same disadvantages as the approaches at the systems level. For instance, Lewen, Bialek, and De Ruyter van Steveninck (2001) mention that the neural response is extremely biased when using open-loop measuring techniques (in neural studies the animals are usually fixated or dead; see also Egelhaaf *et al.*, 2002). In addition, Stanley (2002) argues that the receptive field properties of visual neurons in natural conditions vary over time, because of adaptations to environmental change. The variations cannot be observed in traditional neural approaches that assume time-invariant input-output patterns. Efforts have been started to record neural activity in animals that are stimulated by natural spatio-temporal input patterns (Kern, Petereit, and Egelhaaf, 2001b; Egelhaaf *et al.*, 2001). However, so far it remains difficult to apply neural approaches to freely moving animals.

Another disadvantage of neural approaches is due to the wide variety of neural mechanisms in nature. For instance, Fritzsche (1995) shows that the neural oculomotor system of humans and sharks evolved from that of lampreys in different ways at the neural level. Therefore, neural mechanisms are rather specific. Determining the generic causes of the trade-off in a set of rather different animals requires the use of generalised models of visually guided behaviour. The construction of such generalised neural models appears feasible only at the early stages of visual processing (cf. Barlow, 1961; Van Hateren, 1992a).

Altogether, we conclude that neural approaches and the corresponding level of modelling provide valuable insights into the visual resolution of neural systems. However, neural models are rather specific to the animal species investigated. The neural level of modelling is only suitable to study the trade-off in the early stages of visual processing.

2.2.4 The cellular level

At the cellular level, the physical and chemical properties of individual visual neurons are investigated. The level comprises the approaches termed cellular studies (using animals) and cellular modelling.

Cellular studies

Cellular studies (Revest and Longstaff, 1998) are used, *e.g.*, to explain the spatio-temporal resolution of individual neurons by their cellular properties (membrane properties, molecular composition, etc.). For instance, Marcos *et al.* (1997) determined photopigment densities of photoreceptors in the human eye to estimate the spatial resolution.

Cellular modelling

The temporal properties of visual neurons are understood by modelling the molecular processes in the cells (*e.g.*, Jansonius, 1990) as electrical circuits (Kuffler, Nichols, and Martin, 1984). In these models, the neural membranes and the phototransduction mechanisms that control the flow of ions in biological systems (Weckström, Hardie, and Laughlin, 1991) are represented by resistors and capacitances that control the flow of electrons in the circuits (Softky and Koch, 1995). The parameters of the electrical components determine the temporal resolution of the model (Oppenheim *et al.*, 1997). In this way, *e.g.*, De Lange *et al.* (1998) showed how the molecular properties of a photopigment determine the temporal resolution of photoreceptors.

Advantages and disadvantages

All cellular studies investigate visual resolution at the lowest level of our taxonomy, *i.e.*, they yield a rather precise understanding of visual resolution, which can be considered as an advantage. However, cellular studies have two disadvantages as well. First, the studies are usually constrained to specific cell types in specific animals (see, *e.g.*, Kaneko and Shimazaki, 1976; Hamdorf, 1979). Therefore, the results from cellular studies can only be generalised to other animals in limited cases. An example of such a case is the photoreceptor response that applies to many different eye types (Softky and Koch, 1995; Fernald, 2000). Otherwise, generalisation of cellular approaches is troublesome.

Second, the understanding of isolated neural cells does not necessarily explain the behaviour of animals as a whole. The behaviour of complete neural control mechanisms does not only depend on the properties of the constituent neurons, but on the interactions between them as well (Kuffler *et al.*, 1984).

In conclusion, the examples above suggest that cellular studies do not include an appropriate approach to studying the resolution trade-off. However, when the cellular characteristics of visual neurons appear universal, they may be used to construct generalised models.

2.3 Two classes of research

The second dimension in our taxonomy (section 2.1) is the class of research, *i.e.*, animal research and modelling research. In section 2.2 we provided several examples of these classes in relation to the level discussed. Below we present the advantages and disadvantages of either class (subsections 2.3.1 and 2.3.2).

2.3.1 Animal research

Obviously, animal research forms the fundament of all studies of visually guided behaviour. For instance, in chapter 1 we mentioned several examples of this class that have led us to start our research. Therefore, the main advantage of animal research is the fact that they provide the basic data to investigate visually guided behaviour.

However, applying animal research to the study of the trade-off has three disadvantages. First, the many different types of investigation in animal research will lead to results that cannot easily be compared to each other. For instance, Laughlin (1981) mentioned four ways of measuring the temporal resolution of animals. Every way yielded a different value. Consequently, obtaining a dataset of comparable resolution values of animals to study the trade-off is difficult. This is a clear disadvantage of animal research. Second, the limited availability of relevant data is another disadvantage of animal research. Experiments with animals are time-consuming and require much experience. Compiling a large dataset of resolution values is a tedious and laborious work. Therefore, statistically significant analysis of the trade-off using animal data is unfeasible.

Third, we mention as a disadvantage the provision of insight into the control of animal behaviour. Although animal research usually describes the behaviour of biological systems (animals, neural systems, cells) properly, it generally does not provide an explanation of why the behaviour is shown. Explaining the behaviour requires additional experiments (usually at a lower level than before) or the use of models (Egelhaaf and Borst, 1993a). Therefore, the limited provision of insight is a third disadvantage of animal research.

In summary, animal research may provide the raw materials for studying visually guided behaviour. However, we expect that three disadvantages form real obstacles for an adequate study of the trade-off.

2.3.2 Modelling research

Modelling research is commonly used to solve some of the disadvantages of animal research (see Arbib, 1987): first, the use of a single model yields comparable results for different experiments; second, generalised models of biological behaviour yield large quantities of data for statistically significant analysis; and third, models explain the underlying reasons of a particular biological behaviour (see also chapter 6).

Obviously, the use of models may also introduce pitfalls. For instance, Webb (2001) identified the cases in which models fail to produce valuable results. Two examples are: the model is overgeneralised and the model has no biological relevance. To avoid these pitfalls of modelling research, careful evaluation of the models and the results is required (see also chapter 7).

In summary, owing to the advantages and disadvantages mentioned, we assume that using modelling research to studying visually guided behaviour will lead to a satisfactory answer to our research question.

2.4 Two types of investigation

The third dimension in our taxonomy is the type of investigation. For a discussion of the various types of investigation, we refer to the descriptions of animal research and modelling research in section 2.2. In section 2.3 we assumed that modelling research is most suitable for our aims. To determine the best type of modelling investigations, we discuss the advantages and disadvantages of two important types (Pfeifer and Scheier, 1999; Baddeley, Hancock, and Földiák, 2000). They are information theory (subsection 2.4.1) and situated modelling (subsection 2.4.2).

2.4.1 Information theory

Information theory is a widely accepted modelling approach that can be applied at all four levels in the taxonomy. The approach is used to determine the optimal design of control systems such as the visual system (Shannon and Weaver, 1949; Baddeley *et al.*, 2000). Below, we investigate the suitability of the approach to study the resolution trade-off in visually guided behaviour. We start with a brief introduction to information theory. It is followed by some examples that indicate the suitability of the approach.

Introduction to information theory

Information theory views a control system as an information channel (Baddeley, 2000). The channel is characterised by its information capacity C which is expressed as follows (Burton, 2000).

$$C = \frac{1}{2} \int_0^\infty \log_2 \left(1 + \frac{S(\nu)}{N(\nu)} \right) d\nu \quad (2.1)$$

In the equation, ν is used as the temporal frequency, $S(\nu)$ represents the power spectrum of the channel's output signal, and $N(\nu)$ stands for the noise power. Information theory holds the fundamental assumption that a maximal information capacity C yields optimal systems design, *i.e.*, the design is optimal when the noise power, relative to the signal power, of the channel's output signal is lowest. The approach has thus been successfully applied to understand the design of biological

control systems (Barlow, 1961; Van Hateren, 1992b; Laughlin, 1996). For instance, it is understood that filters in the early stages of visual processing represent optimal information channels that minimise the signal-to-noise ratio (SNR) of and redundancy in visual input signals.

Suitability of information theory

Baddeley (2000) indicated two pitfalls in which the application of information theory leads to misunderstanding of biological systems design. First, the information capacity of some neural structures in the brain may not be maximal for reasons of economisation (*e.g.*, improving the SNR requires energy; Laughlin, 2001). As was mentioned above, information theory equates optimal brain design to a maximal information capacity. Therefore, the example presented indicates a pitfall of information theory.

To point at a second pitfall, we follow Baddeley (2000) in stating that the maximisation of information does not always benefit an animal's behaviour. For instance, Postma, Van den Herik, and Hudson (1996) compared the performance of a biological model of visual identification for a range of noise levels. They found that the presence of a low noise level enhanced the performance of the model when compared to a noise-free set-up. Information theory minimises the noise level which, therefore, leads to misunderstanding of the optimal design of this model. Another example of this pitfall, based on our results, will be demonstrated in chapter 6.

In conclusion, based on the pitfalls presented here, we expect that information theory is not the best approach to answer our research question.

2.4.2 Situated modelling

A second main type of modelling investigations that has recently gained attention is the approach of situated modelling (Pfeifer and Scheier, 1998; Webb, 2002). Below, we provide an introduction to the approach, followed by the determination of the suitability of situated modelling for our aims.

Introduction to situated modelling

The approach of situated modelling emphasises the importance of the visual ecology (natural behaviour in natural environments) for the design of biological visual systems. The basic assumption of the approach is that biological sensori-motor systems are designed to produce behaviour that is optimally adapted to their ecology (Clark, 1996; Kortmann, 2001).

To distinguish the approach from other approaches, we mention two features of situated models: (1) situatedness — situated models perceive a realistic environment through their own sensory systems and act upon that environment; (2) embodiment — situated models possess a (physical) body, *i.e.*, they possess sensory systems, control systems, and motor systems. Owing to these features, situated

models differ from traditional models in two ways. First, traditional models usually do not perceive the world through their sensory system. For instance, the world around traditional models may be represented by an optic flow field instead of by light intensities (see, *e.g.*, Pack, Grossberg, and Mingolla, 2001). Second, traditional models often address only parts of sensori-motor systems, such as the retina (*cf.* Hennig and Funke, 2001). Therefore, situated modelling differs significantly from traditional modelling research.

Suitability of situated modelling

The suitability of situated modelling was determined for scientific disciplines such as invertebrate neuroscience (see Webb, 2002 for a survey), experimental psychology (Krichmar and Edelman, 2002), and the philosophy of mind (Varela, Thompson, and Rosch, 1991; Clancey, 1997). For instance, Franceschini, Pichon, and Blanes (1992) studied robot models that emulated the visually guided behaviour of flies to understand why the spatial resolution in the fly eye depends on the centricity of the visual units. Moreover, situated modelling can be applied at all four levels of detail in our taxonomy (Zanker and Zeil, 2001a; Krichmar and Edelman, 2002). Therefore, this type of modelling is quite suitable to study the trade-off.

To determine the suitability more precisely, we mention the following advantage. In contrast to traditional modelling research, situated modelling emphasises the importance of sensori-motor coordination for the behaviour of animals (Pfeifer and Scheier, 1998). Sensori-motor coordination uses visually guided behaviour to structure the input of the visual system actively (see also Ballard, 1991). For instance, inspired by a sensori-motor coordination (*i.e.*, active scanning) found in flies, Violett and Franceschini (1999) presented a situated model that shows tracking and gaze-stabilisation behaviour. Addressing sensori-motor coordination provides a better understanding of visually guided behaviour and the underlying control systems than traditional models (Clark, 1996).

In summary, we discussed the suitability of situated modelling for our research by showing its application in different scientific domains and by indicating its advantages over traditional modelling approaches. Our discussion confirmed that situated modelling may be a suitable approach to answer our research question.

2.5 Situated systems modelling

We chose situated systems modelling, augmented by neural and cellular modelling, as our approach to studying visually guided behaviour. The choice is motivated below by distilling our expectations from the advantages and disadvantages of different approaches identified in the sections 2.2 to 2.4. Moreover, the two main alternative approaches of situated systems modelling are determined for future comparisons (chapter 6).

When comparing the advantages and disadvantages of the four levels identified in subsections 2.2.1 to 2.2.4, we established that the systems level showed sufficient

detail to determine the resolution of visually guided behaviour. Moreover, the systems level allows the comparison of resolution values found in many different animal species. Therefore, we presume that the systems level is appropriate for studying the trade-off. In addition, we established that the neural and cellular level may provide valuable insights into the trade-off in the early stages of visual processing. Therefore, we assume that augmenting the systems level by the neural and cellular levels, when generally applicable, is appropriate when investigating the trade-off.

Moreover, from section 2.3, we understand that, modelling research is more appropriate for our aims than animal research. The former class of research yields more consistent results, is less time-consuming, and has greater explanatory power than the latter. Therefore, of the two classes in figure 2.1 we chose systems modelling, instead of controlled behavioural studies.

Finally, section 2.4 led us to the expectation that using situated modelling is more suitable to studying the resolution of visually guided behaviour than the application of information theory or other traditional modelling types. Therefore, situated modelling was preferred over information theory.

In conclusion, we state that the approach situated systems modelling has two main alternatives. (1) An alternative with respect to the class of research, viz. controlled behavioural studies, and (2) an alternative with respect to the type of investigation, viz. information theory. The summary of advantages and disadvantages above results in our choice for situated systems modelling.

Chapter 3

Scrutinising the trade-off

“As men busied themselves about their various concerns they were scrutinised and studied, perhaps almost as narrowly as a man with a microscope might scrutinise the transient creatures that swarm and multiply in a drop of water.”

*From ‘The war of the worlds’ (p. 1)
by H.G. Wells (1866 – 1946), British writer.*

Chapter contents: Research method; position-detection systems; motion-detection systems; determinants of the trade-off

Having set ourselves the task to use the approach of situated systems modelling to find generic causes of the trade-off, we now face the prevailing question: which research method employs our approach to find the generic causes? To our knowledge, no such method has been conceived so far. Therefore, we developed an own method to answer the research question.

This chapter deals with an exploratory scrutinising of the trade-off, *i.e.*, the chapter serves to establish a method that allows in-depth analyses of the trade-off. In subsection 2.4.2 we decided to investigate situated models of biological systems that control visually guided behaviour. Two biological control systems are selected and explained in section 3.1. To guide our research, we propose five candidates for the generic causes of the trade-off in section 3.2. Our research method is presented in section 3.3.

3.1 Control of visually guided behaviour

There exists a great number of control systems for visually guided behaviour (see, *e.g.*, Arbib, 1987; Alcock, 1993; Kalaska, 1995; Cliff, 1995). Obviously, modelling all systems to find the causes of the trade-off is unfeasible. In contrast, subsection 3.1.1 argues that we may constrain ourselves to modelling the two principal control systems: position-detection systems and motion-detection systems. To make the reader acquainted with each system, they are explicated in subsections 3.1.2 and 3.1.3, respectively.

3.1.1 Two principal control systems

We argue that most occurrences of visually guided behaviour in animals are controlled by two principal systems: position-detection systems and motion-detection systems. We provide four supportive findings for this dichotomy.

First, in the primate brain, Livingstone and Hubel (1988) identified two pathways that connect the eye and the visual cortex: (1) the parvocellular pathway responds to colour and detects visual detail, such as the position of an object; (2) the magnocellular pathway is sensitive to visual motion. Second, in the insect brain, Wehner (1981, p. 422) determined two stimulus parameters, position and velocity, as the decisive input parameters for visually guided tracking and fixation (see also Collett, 1980; Kirschfeld, 1994). Third, Land (1992) found position-detection systems and motion-detection systems in the control of tracking and pursuit behaviour in the primate brain as well as the insect brain. Fourth, Kalaska (1995) suggested that, besides tracking and fixation, the two systems control additional behaviours, such as visually guided grasping, as well.

From the four findings presented above, we assume that position-detection and motion-detection system constitute the principal control systems of visually guided behaviour. A verification of this assumption follows in chapter 7. At this point, we decide to study the trade-off in two situated models that employ either position-detection or motion-detection control systems. The two systems are explicated below.

3.1.2 Position-detection systems

Position-detection systems are used to control visually guided behaviours such as target tracking (see, *e.g.*, Collett, 1980). Target tracking is commonly found in many animals (Land, 1992). Therefore, the behaviour is interesting to study the trade-off in, *e.g.*, insects as well as vertebrates (section 1.2). Below we provide the relevant data from the literature of biological position-detection systems that will be used to construct a situated model of target tracking (chapter 4).

Biological position-detection systems

Biological position-detection systems have been studied at various levels of detail in our taxonomy (section 2.2). For a full survey we refer to the work of Wehner (1981). Here, we constrain ourselves to providing two examples of studies. For instance, the target-tracking behaviour of arthropods and humans was described at the phenomenological level by a proportional control diagram (Land and Collett, 1974; Land, 1992). The tracking response was found to equal a linear fraction of the angle at which the target is detected (Collett, 1980). Therefore, we use a proportional control diagram as a starting point for our model of target tracking in chapter 4. At the neural level, units such as local jittery movement detectors (Wiersma and Yamaguchi, 1967) and on-off cells (Arnett, 1972) are thought to represent parts of position-detection systems at the neural level. Abstract versions of these neural elements will also be used in our model.

3.1.3 Motion-detection systems

Motion-detection systems are fundamental to visual detection (Nakayama, 1985; Zanker and Zeil, 2001b). A prominent type of visually guided behaviour controlled by motion-detection systems is gaze-stabilisation behaviour (Miles and Wallman, 1993). This behaviour, which is displayed by most animals, is described by compensatory movements of the eyes, head, or body to keep a visual scene stable on the retina. Below we survey the current models of motion-detection systems. From the survey we select one to construct a situated model of gaze stabilisation in chapter 5.

Current models of motion-detection systems

All models of motion-detection systems are divided into ‘feature-tracking’ and ‘intensity-based’ models (Seiffert and Cavanagh, 1999; Harrison and Koch, 2000). Feature-tracking models (cf. Ullman, 1983) calculate the time needed by an identified object to travel a given distance. The fraction of the distance over the time equals the velocity of the detected motion. However, feature-tracking models are less biologically plausible, since their behaviour is not robust in noisy circumstances (Harrison and Koch, 2000). In contrast, intensity-based models are plausible since their response was shown to resemble the response of biological motion-detection systems (Borst and Egelhaaf, 1989). Intensity-based models of motion-detection systems are further divided into gradient models and correlation-based models. Correlation-based models are the most generally accepted models of motion-detection and, therefore, are discussed below.

Models of correlation-based motion detectors include energy models (Adelson and Bergen, 1985), elaborated Reichardt detectors (Van Santen and Sperling, 1985), and the Barlow-Levick model (Barlow and Levick, 1965). Although these models all emulate biological motion detection (Borst and Egelhaaf, 1989), the dominant correlation-based model is the Reichardt correlator (Hassenstein and Reichardt, 1956; Reichardt, 1961). The basic Reichardt correlator was already discussed in

chapter 2 (figure 2.2.2) and comprises the principal elements needed to perform motion detection: two input channels, a non-linear interaction (*e.g.*, multiplication), and an asymmetry (Borst and Egelhaaf, 1989).

In most studies, the basic Reichardt correlator is adapted in two ways to increase the biological plausibility of the correlator (figure 3.1.3). First, the linear delay term

Figure 3.1: The symmetric Reichardt correlator model for biological motion detection.

Δt is replaced by temporal neural filters. For instance, Mastebroek *et al.* (1980) used a combination of two first order filters (an integrating, low-pass filter *I* and a differentiating, high-pass filter *D*). Moreover, most studies employ symmetric Reichardt correlators. As is shown in figure 3.1.3, the output of a correlator that is sensitive to leftward motion (the circle labelled 'X' on the right of the diagram) is subtracted from the output of a rightward sensitive correlator (circle labelled 'X' on the left of the diagram). The symmetric correlator is rightward sensitive. It was determined that symmetric correlators resemble the behaviour of biological motion detectors better than non-symmetric correlators do, *i.e.*, they are less sensitive to changes in overall light intensity and more directionally sensitive than non-symmetric ones (Borst and Egelhaaf, 1990). Owing to this determination, we employed symmetric Reichardt correlators in our model of gaze-stabilisation (chapter 5).

To conclude the section, we state that finding the generic causes of the trade-off requires models of the principal control systems of visually guided behaviour. Based on what was determined above, we decided to pursue the following two modelling studies. In chapter 4 we construct and investigate a model of target tracking based on a position-detection control system. Chapter 5 addresses the construction and investigation of a model of gaze-stabilisation based on a motion-detection system.

3.2 Candidate causes of the trade-off

In chapter 1 we formulated our research question as “which are the generic causes of the trade-off between spatial and temporal resolution in visually guided behaviour?”

Below, we determine candidate causes, called ‘determinants of the trade-off’. In subsection 3.2.1 the term determinants is explained and five of them are identified. Thereafter, in subsections 3.2.2 to 3.2.6, each determinant is discussed in detail.

3.2.1 Five determinants of the trade-off

We define a determinant of the trade-off as a factor that possesses the following three properties: (1) the factor is significant for evolutionary success; (2) the factor depends on the spatial and temporal resolution; (3) optimising the factor to accomplish successful behaviour may cause a trade-off between spatial and temporal resolution.

We identify five determinants as follows. Barlow (1981) identified three reasons why the spatial resolution of the human eye is lower in the periphery than in the fovea: (1) to enhance sensitivity; (2) to economise central processing; and (3) to improve depth estimation from stereovision. From the first reason (enhancing sensitivity) we assume that the factors light sensitivity, contrast sensitivity, and motion sensitivity are determinants. The assumption is motivated in subsections 3.2.2 to 3.2.4. The second reason (economising central processing) yields the determinants heat production and energy consumption, as argued in the respective subsections 3.2.5 and 3.2.6. We do not consider Barlow’s third reason since stereoscopic depth perception is not used by all animals in the world. For instance, small insects such as honeybees (Lehrer, 1997) or locusts (Sobel, 1990) use *active* depth perception instead, due to the relatively small inter-ocular distance that weakens stereoscopy.

3.2.2 Light sensitivity

We show that the factor light sensitivity is a determinant by pointing at the three properties mentioned in subsection 3.2.1.

Significance for evolutionary success

Light sensitivity determines the quality of visual detection (Wandell, 1995). It is generally accepted that good visual detection provides a selective advantage in evolution (see, *e.g.*, Fernald, 2000). Therefore, we derive that light sensitivity is significant for evolutionary success.

Dependency on spatial and temporal resolution

The number of photons absorbed by a photoreceptor is a measure for light sensitivity (Land, 1981). Below, we establish the dependency of this number (and therefore of the factor light sensitivity) on the spatial and temporal resolution (cf. Kortmann, Postma, and Van den Herik, 1999; Warrant, 1999). Foremost, we show this fact for compound eyes, commonly found in insects. Thereafter, we generalise it to apply to other types of eyes as well.

Compound eyes Compound eyes contain a mosaic of visual units called ommatidia. Each ommatidium possesses a facet lens through which light is projected on a number of photoreceptors (Snyder, 1979; Land, 1997). The spatial resolution of the eye is determined by the angular distance between neighbouring ommatidia — the interommatidial angle $\Delta\phi$ — and by the angular receptance angle $\Delta\varrho$, *i.e.*, the maximal angle at which incident light is detected (Snyder, 1979). The temporal resolution can be measured in different ways (Laughlin, 1981). We employ the integration time Δt which is defined as “the largest time interval within which all incident light is treated [...] as if it arrived at the same instant.” (Laughlin, 1981, p. 223).

The dependency of light sensitivity on spatial and temporal resolution is determined as follows. Snyder (1979) provides a formula for the mean number of photons caught by photoreceptors in a compound eye per integration time. Without loss of generality, we simplify Snyder’s formula to express the photon catch of a 1-dimensional compound eye (cf. chapter 4).

$$\bar{N} = 0.89 \cdot \epsilon \cdot \Delta t \cdot \bar{I} \cdot D \cdot \Delta\varrho \quad (3.1)$$

In this formula, $\bar{N}$ indicates the average number of photons; ϵ represents the quantum efficiency (constant); Δt is the integration time; $\bar{I}$ equals the mean number of photons captured per square μm , per steradian of field (proportional to stimulus luminance L : $\bar{I} = k_1 \cdot L$); D denotes the facet diameter (proportional to the interommatidial angle: $D = \eta_1 \cdot \Delta\phi$); and $\Delta\varrho$ expresses the angular receptance angle (also proportional to the interommatidial angle: $\Delta\varrho = \eta_2 \cdot \Delta\phi$).

Equation (3.1) is reformulated to express the light sensitivity (S_L) as a function of spatial and temporal resolution (Δt and $\Delta\phi$, respectively).

$$\begin{aligned} S_L = k_2 \cdot \bar{N} &= k_2 \cdot 0.89 \cdot \epsilon \cdot \Delta t \cdot k_1 \cdot L \cdot \eta_1 \cdot \Delta\phi \cdot \eta_2 \cdot \Delta\phi \\ &= k_3 \cdot \Delta t \cdot (\Delta\phi)^2 \end{aligned} \quad (3.2)$$

with k_2 as a constant determining the number of photons needed for reliable detection and $k_3 = k_1 \cdot L \cdot k_2 \cdot 0.89 \cdot \epsilon \cdot \eta_1 \cdot \eta_2$. Equation (3.2) shows that the light sensitivity of compound eyes depends on the spatial and temporal resolution.

Other types of eyes The dependency can be generalised to other types of eyes. For instance, in single-lens eyes (commonly found in vertebrate species) the incoming flux of photons is divided over the number of photoreceptors in the retina. Assuming a maximal absorption of the flux, a high density of photoreceptors leads to a low photon catch per receptor. Therefore, high spatial resolution yields low light sensitivity.

Moreover, a short integration time results in a low number of photons caught, in either type of eye. Therefore, high temporal resolution yields a low light sensitivity.

Altogether, light sensitivity depends on the spatial and temporal resolution in both compound and single-lens eyes (cf. Warrant, 1999).

Light sensitivity is a candidate cause of the trade-off

To determine that the factor light sensitivity is a candidate cause of the resolution trade-off we show that for a fixed light sensitivity, the variables $\Delta\phi$ and Δt lie on a trade-off curve.

When substituting $\frac{S_L}{k_3} = k^2$ we obtain equation (3.3).

$$\Delta\phi = \frac{k}{\sqrt{\Delta t}} \quad (3.3)$$

The equation describes a trade-off curve, scaled by the value of k as is depicted in figure 3.2. Therefore, we conclude that the factor light sensitivity is a candidate cause of the trade-off.

Figure 3.2: A trade-off curve caused by light sensitivity and scaled by the factor k . The axes have a linear scale and denote arbitrary units.

3.2.3 Contrast sensitivity

Below we argue that the factor contrast sensitivity is the second determinant of the trade-off.

Significance for evolutionary success

Comparably to light sensitivity, contrast sensitivity determines the quality of visual detection (Wandell, 1995). Therefore, the factor is significant for evolutionary success.

Dependency on spatial and temporal resolution

Sjöstrand (2001) determined that the human eye has dedicated mechanisms that enhance contrast to improve spatial resolution. He concluded that “contrast enhancement is therefore a major factor contributing to visual acuity” (p. 1). Therefore, we suggest that contrast sensitivity and spatial resolution are somehow related. Below, we show that contrast sensitivity depends on spatial and temporal resolution.

Contrast sensitivity CS is defined as the inverse value of the contrast threshold ϑ_C required for reliable visual detection: $CS = 1/\vartheta_C$ (Wandell, 1995). Whether the visual detection is reliable depends mainly on the SNR of the photoreceptor response (Barlow, 1956). The highest SNR yields the most reliable detection.

The SNR of a photoreceptor’s response to a moving sine-wave grating of spatial frequency f_s , velocity v , mean intensity I , and modulation ΔI is defined below (Snyder, 1979; Warrant, 1999).

$$SNR = \sqrt{\bar{N}} \cdot c \cdot M(f_s, f_t) \quad (3.4)$$

with $\bar{N}$ representing the mean photon catch (subsection 3.2.2), $c = \Delta I/I$ the contrast of the grating, $f_t = v \cdot f_s$ the temporal grating frequency and $M(f_s, f_t)$ the photoreceptor’s spatio-temporal amplitude response function.

Using a fixed contrast c , Warrant (1999) found that decreasing the spatial and temporal resolution enhances the SNR, *i.e.*, a decrease of visual resolution will increase the product $\sqrt{\bar{N}} \cdot M(f_s, f_t)$ in equation (3.4). Using this finding we derive that for decreasing values of c , a reliable detection is maintained when $\bar{N}$ is increased, *i.e.*, when the resolution is decreased. Therefore, we determine that the contrast sensitivity CS depends on the spatial and temporal resolution.

Contrast sensitivity is a candidate cause of the trade-off

We show why the factor contrast sensitivity is a candidate cause of the trade-off. For this we first use equation (3.4) to express the contrast sensitivity CS in terms of the SNR threshold ϑ_{SNR} that is required for reliable visual detection.

$$CS = \frac{1}{\vartheta_C} = \frac{\sqrt{\bar{N}} \cdot M(f_s, f_t)}{\vartheta_{SNR}} \quad (3.5)$$

Having done so, we find in equation (3.5) that the contrast sensitivity depends on the square root of the photon catch $\bar{N}$. In subsection 3.2.2 it was determined that the $\bar{N}$ is a measure of the factor light sensitivity. This factor causes that spatial and temporal resolution values lie on a trade-off curve (figure 3.2). Therefore, from equation (3.5), we expect that contrast sensitivity causes the existence of similar curves, *i.e.*, we suspect that the factor contrast sensitivity causes a trade-off between spatial and temporal resolution.

3.2.4 Motion sensitivity

In subsection 3.2.1, we identified the factor motion sensitivity as the third determinant of the resolution trade-off.

Significance for evolutionary success

As was mentioned in section 1.1, the timely detection of moving objects is important to, *e.g.*, evade predators, or successfully chase mates for reproduction. Therefore, motion sensitivity is significant for evolutionary success.

Dependency on spatial and temporal resolution

Obviously, motion sensitivity benefits from high temporal resolution, since fast moving objects are detected quickly. In addition, also the spatial resolution sets the motion sensitivity (Kortmann, Postma, and Van den Herik, 2000). Tiny movements are more rapidly detected when a high spatial resolution is employed than when using low resolution. In conclusion, the factor motion sensitivity depends on the spatial and temporal resolution.

Motion sensitivity is a candidate cause of the trade-off

Srinivasan and Bernard (1975) have shown that the acuity of a visual system is determined by the product of spatial and temporal resolution, as well as by the velocity of a moving target. Fast moving targets cause blurred retinal images, which can be de-blurred by increasing the product of spatial and temporal resolution (see also Laughlin, 1981; Warrant, 1999). Therefore, when defining motion sensitivity MS as the target velocity that causes a threshold level of blurring, MS is calculated as a function $\mathcal{F}$ of the product of spatial and temporal resolution.

$$MS = \mathcal{F}(\text{spatial resolution} \times \text{temporal resolution}) \quad (3.6)$$

The equation indicates that the optimisation of motion sensitivity is achieved by maximising the product of spatial and temporal resolution. This finding contrasts the way in which the factors light sensitivity and contrast sensitivity are optimised (*viz.* by minimising the resolution). As a result, we expect that an increase of spatial and temporal resolution must be compensated for by a decrease of temporal

resolution, and *vice versa* (i.e., the optimal resolution values lie on a trade-off curve as the one shown in figure 3.2; see also Kortmann *et al.*, 1999). Therefore, we assume that the factor motion sensitivity causes a trade-off between spatial and temporal resolution.

3.2.5 Heat production

We discuss why the factor heat production is our fourth determinant of the trade-off.

Significance for evolutionary success

It has been argued that the evolutionary development of the brain to the extent as found in humans was limited by the levels of heat production and heat dissipation (Baker, 1979; Falk, 1992). The metabolic activity in photoreceptors and neurons (e.g., the operation of the ion pumps in the cell membrane) produces heat in the neural system. The neurons are cooled by a heat-dissipation system. In fact, different dedicated heat-dissipation systems have been identified which prevent the brain temperature from reaching lethal levels (Baker, 1979).

For instance, to dissipate heat dogs possess a heat-exchange network in the inner brain called a *rete* (Baker and Chapman, 1977). This type of heat-dissipation systems has a relatively low capacity as compared to the cranial-radiator systems found in the higher primates (Zenker and Kubik, 1996). Cranial radiators consist of small blood vessels that traverse the skull (*cranium*) in the fashion of a radiator. The blood in the radiator travels from the inside to the outside of the skull and returns inward towards the brain after exchanging heat with the skin. The latter type of heat-dissipation system provides a much better cooling capacity than the *rete*-based systems found in lower species, such as dogs. Therefore, the human brain can produce more heat than the brain of dogs.

One of the distinctive features between the human brain and that of lower (vertebrate) species is the ratio of brain mass over body mass. The brains of humans are relatively heavier than those of lower mammals, such as dogs (Jerison, 1985). When comparing relative brain mass of (hot-blooded, terrestrial) animals to the capacity of the heat-dissipation system, we find a positive correlation (Falk, 1992). For instance, the relative brain mass of primates (high-capacity, cranial-radiator type dissipation system) is larger than that of dogs (low-capacity, *rete*-based system). Moreover, evolutionary studies of fossil hominid skulls report that relative brain mass is correlated with the capacity of the cranial radiator (Falk and Gage, 1998). Therefore, it is likely that relative brain mass is limited by the capacity of the heat-dissipation system. In conclusion, the factor heat production is a significant factor for evolutionary success.

Dependency on spatial and temporal resolution

We first discuss the dependency of heat production on spatial resolution, followed by temporal resolution.

Dependency on spatial resolution To understand why heat production depends on the spatial resolution, we remark that high spatial resolution in the eye generally yields a large number of photoreceptors. As a result, to preserve acuity in the brain areas that process visual stimuli, high spatial resolution demands large central brain areas. We assume that the number of photoreceptors does not exceed the size of central brain areas dedicated to vision. The assumption is motivated by the fact that the total number of photoreceptors in a fly's eye (40,000; Beersma, Stavenga, and Kuiper, 1977) roughly equals the number of neurons in the highest visual centre of the fly brain (lobula plate; 39,600 neurons on each side of the brain; Strausfeld, 1976). Moreover, the number of photoreceptors in the human retina (approximately 100 million) is smaller than the number of neurons in the primary visual cortex V1 (ca. 2000 million per hemisphere, Van Essen and Anderson, 1990). As is argued above, large brains generate more heat than small brains and therefore need more cooling. Therefore, our discussion suggests that the heat production of the brain depends on the spatial resolution of the visual system.

Dependency on temporal resolution In congruence to the case of increasing spatial resolution, we suggest that enlarging the temporal resolution of a neuron results in more heat production. The increased heat production is due to higher metabolic activity of the ion pumps (Laughlin, 1996). As with spatial resolution, we assume a linear, or higher-order, relationship between temporal resolution in the retina and in the central areas of the brain (cf. Zaagman, Mastebroek, and De Ruyter van Steveninck, 1983; Livingstone and Hubel, 1988). Therefore, it is suggested that the heat production in the central brain depends on the temporal resolution in the retina.

Heat production is a candidate cause of the trade-off

Altogether, increasing both spatial and temporal resolution in the retina generally amounts to increased heat production in the central brain (see also Yablonskiy, Ackerman, and Raichle, 2000). A fixed capacity of the heat dissipation system (determined by the type of cooling system, the degree of vascularisation, etc.) therefore sets an upper limit to heat production and thus to the total level of spatial and temporal resolution of a visual system. High levels of temporal resolution need be compensated for by low levels of spatial resolution, *i.e.*, the resolution values lie on a trade-off curve (figure 3.2). This observation suggests that the factor heat production is a candidate cause of the trade-off.

3.2.6 Energy consumption

The fifth and last determinant to be discussed is the energy consumption of the brain.

Significance for evolutionary success

The metabolic costs for photoreception are relatively high as compared to other energy-consuming processes. For instance, Laughlin *et al.* (2000) mention that a blowfly uses up to 10% of its resting energy consumption on photoreceptors. An even higher fraction has been reported for the human brain (20%; Kety, 1957), whereas the brain only accounts for about 2% of the total human body mass. Of the total energy consumption in the mammalian brain, 50 – 80% is spent on signalling, *i.e.*, the operation of the sodium pumps in neurons that produce action potentials (Laughlin, 2001). Therefore, we conclude that the factor energy consumption, as well, is significant for evolutionary success.

Dependency on spatial and temporal resolution

The total number of neurons in the retina and central brain is bounded by energy supply: “the supply of metabolic energy constrains the growth and function of many organs, including sensory organs and the brain” (Laughlin, 2001, p. 475). For a fixed visual angle, the number of photoreceptors, by definition, sets the angular density of the photoreceptors¹. High density results in high spatial resolution (section 1.2.1). Therefore, we argue that the energy consumption depends on the spatial resolution.

The temporal resolution of a visuo-motor system determines the rate of energy consumption (Laughlin, 1996) which, in turn, also sets the requirements for the capacity of the energy-supply system. Therefore, the factor energy consumption also depends on temporal resolution.

Energy consumption is a candidate cause of the trade-off

Given a limited rate of energy supply to the visual system we determine that a high spatial resolution is compensated for by a low temporal resolution to keep the energy consumption within bounds, *i.e.*, the resolution values lie on a trade-off curve (figure 3.2). In conclusion, the factor energy consumption is a candidate cause of the trade-off.

3.3 How to find generic causes

In this chapter we developed a method to scrutinise the trade-off between spatial and temporal resolution in visually guided behaviour. It is used to find generic causes of the trade-off. In section 1.3 we defined generic causes as those causes that (1) explain the existence of a trade-off in many occurrences of visually guided behaviour, (2) cause a trade-off in many different biological visual systems, and (3) cause a trade-off in artificial visual systems.

¹The angular density is defined as the fraction of the number of photoreceptors over the visual angle.

For our method, we selected the two principal control systems of visually guided behaviour: position-detection systems and motion-detection systems (section 3.1). Above it was shown that these two systems control many types of behaviour and that they are found in many animals as well as in artificial systems. Therefore, we expect that the causes of a trade-off in position-detection systems and motion-detection systems are generic causes of the trade-off. Moreover, in section 3.2 we selected five determinants: factors that are significant for evolutionary success, that depend on spatial and temporal resolution and that are expected to cause a trade-off when optimised. The five determinants are: light sensitivity, contrast sensitivity, motion sensitivity, heat production, and energy consumption.

To determine the generic causes we will verify whether the five determinants cause a trade-off between the spatial and temporal resolution of a situated model of a position-detection system (chapter 4) and of a motion-detection system (chapter 5). The verification will happen as follows. We maximise the behavioural success of our models by tuning the spatial and temporal resolution to values that optimise the determinants of the trade-off (optimal light sensitivity, optimal contrast sensitivity, etc.). To determine whether optimisation of the determinants leads to a trade-off, we tune the resolution values for a range of visual ecologies (*i.e.*, for a range of values of typical light intensity, image contrast, velocity of visual motion, heat dissipation, and energy supply). If the tuned resolution values lie on a trade-off curve, we conclude that the determinants cause a trade-off.

Chapter 4

The resolution of position-detection systems

“I don’t think of form as a kind of architecture. The architecture is the result of the forming. It is the kinesthetic and visual sense of position and wholeness that puts the thing into the realm of art.”

Roy Lichtenstein (1923 – 1997), U.S. Pop artist.

Chapter contents: Position-detection systems, visual resolution, light sensitivity and motion sensitivity, heat production, energy consumption, evolutionary algorithms¹.

Position detection is a principal mechanism that controls visually guided behaviour such as target tracking (see subsection 3.1.2 and Collett and Land, 1975; Wehner, 1981; Land, 1992; Kirschfeld, 1994). In this chapter we investigate the resolution of position-detection systems.

4.1 Aim of the chapter

We aim to verify whether four determinants (light sensitivity, motion sensitivity, heat production, and energy consumption; section 3.2) cause a trade-off between the spatial and temporal resolution of a position-detection system. The fifth determinant, contrast sensitivity, is not investigated due to the simplicity of the visual

¹This chapter is an updated version of Kortmann, R. Postma, E.O. and Van den Herik, H.J. (2001). Evolution of visual resolution constrained by a trade-off. *Artificial Life*, Vol. 7, No. 2, pp. 125-145. Thanks are due to the MIT Press for allowing the reuse of the paper.

stimulus used in the experiments that follow: the stimulus is defined as a moving point source of light of which the contrast is determined by the mean intensity. The optimal contrast sensitivity is therefore directly set by the first determinant, light sensitivity. The importance of contrast sensitivity will be studied in chapter 5 where we investigate more complex input stimuli.

4.1.1 Chapter survey

The chapter is organised as follows. In section 4.2 we construct a model of target tracking that is based on a position-detection system. Thereafter, in section 4.3 we show how an evolutionary algorithm can optimise the resolution of the model. In section 4.4, we demonstrate the evolution of a trade-off when optimising the determinants light sensitivity and motion sensitivity. Then we optimise the determinants motion sensitivity and heat production (section 4.5) and the determinants motion sensitivity and energy consumption (section 4.6). In section 4.7, we compare our results to the evolved values of spatial and temporal resolution in biological visual systems. The comparison shows that the determinants are expected to cause a trade-off in our situated model of target tracking, as well as in biological position-detection systems. Finally, we discuss the results and provide chapter conclusions in section 4.8.

4.2 A model of target tracking

In this section we present a model of target tracking by which we investigate the trade-off between spatial and temporal resolution in position-detection systems. The model comprises the basic aspects found in the tracking systems of different animals ranging from invertebrate species (*e.g.*, insects) to the human visual system. We emphasise that the model is not meant to explain every detail of the target tracking behaviour as displayed by animals. Instead, it is used to verify whether our determinants cause a trade-off between spatial and temporal resolution in a model that captures the basic principles of biological position-detection systems. To explain the model, we compare the target-tracking mechanisms of different animal species in subsection 4.2.1. From the comparison, we constructed a model of which we introduce the components in subsection 4.2.2. The model is implemented in a mobile robot simulator as outlined in subsection 4.2.3. In subsection 4.2.4, we describe the model's most important component: the position-detection mechanism. Subsection 4.2.5 summarises the section.

4.2.1 Types of target tracking

Target tracking is the visually guided behaviour in which animals move their eyes, head, or complete body to follow a moving target (see, *e.g.*, Land and Collett, 1974; Lisberger, Morris, and Tychsen, 1987; Kirschfeld, 1994; Krauzlis and Stone,

1999). In most visual systems, only the central parts of the retina have sufficiently high acuity to perform object recognition tasks (Wandell, 1995). Therefore, target tracking is necessary to identify moving objects. Alternative purposes of target tracking are to chase other animals, such as prey or mates (Land and Collett, 1974; Gilbert and Strausfeld, 1991).

There exist different types of tracking systems that are characterised by different types of input and output signals. For instance, humans use information about the position as well as the velocity of a moving target to keep the target fixed in the centre of the visual field (Land, 1992). In contrast, the tracking mechanisms in hoverflies is purely position-driven (Collett and Land, 1975). Both systems produce a smooth pursuit behaviour when the target velocity is relatively small. The tracking behaviour of many animals switches from smooth pursuit to an intermittent pattern involving smooth movements and saccades (high velocity turns) when the target velocity increases or when the background is clearly textured (Kirschfeld, 1994). The saccadic system is driven by both position-detection and motion-detection systems. In our research we constrain ourselves to smooth-pursuit behaviour controlled by a position-detection system, since this type of behaviour and the underlying control system is the most basic type found in biological visual systems (Land, 1992).

4.2.2 Components of the model

Our model of target tracking contains many components. As mentioned before, the model comprises the most basic control system for target tracking, which is a proportional negative feedback controller (figure 4.1; cf. Collett, 1980; Land, 1992). The input of the controller is the angular position of the target, θ_i . In figure 4.2 we

Figure 4.1: The control diagram for target tracking.

illustrated the angular position of the target (θ_i), of the model (θ_o), and the difference between the two angular positions ($\theta_e = |\theta_i - \theta_o|$). In the controller, the input value θ_i is compared to the current angular position of the model, θ_o . The difference between θ_i and θ_o is the position error θ_e . A position-detection mechanism, consisting of a visual sensor and an artificial neural network, transforms the position error into a velocity signal. As in biological systems, the position error θ_e is approximated by a

Figure 4.2: The angular position θ_o of the model of target tracking and the angular position θ_i of the target (*). Both positions are shown with respect to an arbitrary reference direction. Also shown is the position error $\theta_e = |\theta_i - \theta_o|$.

visual sensor as $\theta'_e = \theta_e \pm \epsilon$. In this equation, ϵ is a detection error and depends on the precision of the sensor.

After approximation of the position error, the controller adds noise to yield θ_n , the noisy approximation of the position error. The type of noise is chosen such to emulate the noise in biological visual systems (photon shot noise; Snyder, 1979). Photon shot noise is caused by the random arrival of photons at the retina and is best described by a Poisson distribution (Land, 1981). We approximate the Poisson distribution using a Gaussian with standard deviation equal to the square root of the mean (see subsection 4.2.4 for details). Earlier experiments have shown that this approximation holds well for values close to the mean (highest chance of occurring).

The signal θ_n is subsequently fed into an artificial neural network. We employ an adapted version of a perceptron with a linear transfer function (Hertz, Krogh, and Palmer, 1991) to produce an output signal $\dot{\theta}_o$ (see subsection 4.2.4). This output signal represents the angular velocity of the model, *i.e.*, the motor behaviour is restricted to rotation around the vertical axis. Finally, as is shown in figure 4.1, integrating the output $\dot{\theta}_o$ yields the new angular position θ_o that is fed back in the control loop in the next simulation step.

The control diagram shown in figure 4.1 represents the behaviour of the model which is implemented as described in the following subsection.

4.2.3 Implementation

The model is implemented in a mobile robot simulator² based on the Khepera robot (Mondada, Franzi, and Ienne, 1994). The Khepera robot is a miniature mobile robot shown in figure 4.3. It has two wheels which are driven by separate motors thus enabling the robot to turn around its vertical axis. Moreover, it has several sensors for measuring distances to objects, ambient light intensity, etc.

²The simulator, written by Olivier Michel, is free of charge for educational and research purposes and can be downloaded from the internet at <http://diwww.epfl.ch/lami/team/michel/khep-sim/>

Figure 4.3: The Khepera mobile robot.

In the following sections we optimise the resolution of the model using an evolutionary algorithm. Although it is possible to apply this type of algorithms to optimise the behaviour of a physical robot (Floreano and Mondada, 1996; Kortmann and Hallam, 1999), the optimisation process is significantly accelerated when using a computer-simulated version of the robot. We employ a simulator of which a screenshot is shown in figure 4.4. In the figure, a top view of the robot and its environment is shown on the left. Also shown in the left part of the figure is the

Figure 4.4: Khepera robot simulator by Olivier Michel.

target that is tracked (black spot). The target is defined as a point-shaped light source orbiting the robot at fixed radius. In the right part of the figure, different types of sensor readings are visualised. The simulator models the motor behaviour of the visuo-motor system, *i.e.*, the velocity integration shown in the control diagram. We extend the simulator for our purposes by adding a visual sensor and an artificial neural network that implement the position-detection mechanism (see figure 4.1). This is described in the next subsection.

4.2.4 Position-detection mechanism

We emphasise that we aim at constructing a *situated* model of target tracking (chapter 2). Therefore, the model determines the position error θ_e from its sensory input using a visual sensor. In this subsection we first discuss the visual sensor, followed by the artificial network. Finally, we show how the spatial and temporal resolution are represented in the model.

Visual sensor

The visual sensor resembles the initial stages of visual processing of biological visual systems (Srinivasan and Bernard, 1975; Laughlin, 1981; Warrant, 1999). The optics of the visual sensor are modelled after the compound eyes of insects. We note that the projection of the visual world on the photoreceptors in compound eyes is not fundamentally different from that in single-lens eyes (Kirschfeld, 1976). Therefore, the model is not restricted to insects. Compound eyes consist of multiple ommatidia. The visual sensor consists of a one-dimensional array of simulated ommatidia, shaped as a semi-circle on the rim of the simulated robot body.

Each simulated ommatidium consists of a facet lens that projects light onto exactly one photoreceptor. Equation (4.1) expresses the photoreceptor response, R , as the product of the target's illumination I_T , the interommatidial angle $\Delta\phi$ (the angular distance between adjacent photoreceptors), and a Gaussian function of the angle of incident light α measured as the deviation from the photoreceptor's midline (Kirschfeld, 1969; Snyder, 1979).

$$R = I_T \cdot \Delta\phi \cdot e^{-2.77 \cdot (\frac{\alpha}{\Delta\phi})^2} \quad (4.1)$$

The Gaussian's half-width is represented by $\Delta\phi$, the photoreceptor's acceptance angle. We keep $\Delta\phi$ proportionate to the interommatidial angle $\Delta\phi$ (Snyder, 1979). The factor 2.77 is used for normalisation (with $\alpha = \frac{1}{2}\Delta\phi$ the response R equals $\frac{1}{2} \cdot I_T \cdot \Delta\phi$).

As discussed in subsection 4.2.2, the photoreceptor response is subject to photon shot noise. In the model, the actual photoreceptor response R' is drawn from a Gaussian distribution with mean $\mu = R$ and standard deviation $\sigma = \sqrt{R}$. The normalised chance P of response R' is thus calculated.

$$P(R', \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(R' - \mu)^2}{2\sigma^2}} \quad (4.2)$$

The response array of the receptors in the visual sensor is used as the input pattern θ_n of an artificial neural network that calculates the response velocity $\dot{\theta}_o$ of the system (figure 4.1).

Artificial neural network

The artificial neural network that maps the noisy sensor signal θ_n to the system's angular velocity $\dot{\theta}_o$ is visualised in the lower right part of figure 4.4. In the figure, the array of photoreceptors is shown as a row of small dots (top layer). The photoreceptors project their output onto a retinotopic array of input neurons (small rectangles; one-to-one projection). The input neurons (one for each photoreceptor) integrate their input over a period Δt simulation steps. The variable Δt models the receptor's integration time determined by the phototransduction process (section 3.2.2). Using discrete time steps t , the time-dependent activation $I_i(t)$ of input neuron i is calculated as follows.

$$I_i(t) = \sum_{t'=\Delta t \cdot (t \operatorname{div} \Delta t)}^t T_{\text{step}} \cdot R'_i(t') \quad (4.3)$$

In this equation we use the operator div as the integer division; the variable T_{step} represents the duration of one simulation step; and $R'_i(t')$ is the time-dependent response of a photoreceptor corresponding to input neuron i (equation (4.2)). After each Δt -interval the input neurons engage in a winner-take-all competition, which models the lateral inhibition found in biological visual systems (Fiorentini *et al.*, 1990). Subsequently, all activation is flushed which represents the operation of the ion pumps that returns the activation of the cells to the resting level (Laughlin, 1996). After the winner-take-all competition, the winning input neuron represents the direction of the target light source, relative to the system's heading. The activation of this neuron is propagated to a single output neuron by multiplying the activation by a weight value w .

To implement a proportional controller, the weight values w_i of the connection between input neuron i and the output neuron increase linearly with angular distance ϕ_i of input neuron i to the central input neuron (see also Wehner, 1981).

$$w_i = -\frac{2 \cdot \phi_i}{\pi} \quad \text{if } i \in \{1, 2, \dots, \frac{N}{2}\} \quad (4.4)$$

$$w_i = \frac{2 \cdot \phi_i}{\pi} \quad \text{if } i \in \{\frac{N}{2} + 1, \frac{N}{2} + 2, \dots, N\} \quad (4.5)$$

using N as the number of photoreceptors. The activation of the output neuron is interpreted as the output signal $\dot{\theta}_o$, *i.e.*, the angular velocity, of the system.

All neurons have a linear activation function. As a result, the output signal is proportional to the eccentricity of the winning neuron. This fact reflects the target tracking mechanisms in biological systems. For instance, the hoverfly *Syrretta*

pipiens L. increases the steering signal proportionally to the angle at which the target is perceived (e.g., Collett and Land, 1975).

We recall that the model introduced above is used to determine whether the determinants cause a trade-off between spatial and temporal resolution in position-detection systems. In order to do so, we indicate in the following paragraph how the spatial and temporal resolution are represented in the model.

Spatial and temporal resolution

In the model of target tracking the spatial and temporal resolution are represented as follows. The interommatidial angle $\Delta\phi$ determines the receptor density and thus the spatial resolution. High values of the angle $\Delta\phi$ indicate low spatial resolution (low receptor density) and *vice versa*. Since the array of photoreceptors covers a semi-circle (π rad) we use $\Delta\phi$ to calculate the number of photoreceptors N_{sensors} .

$$N_{\text{sensors}} = \frac{\pi}{\Delta\phi} + 1 \quad (4.6)$$

It is remarked that sometimes the photoreceptor acceptance angle $\Delta\theta$, in addition to $\Delta\phi$, is used to define spatial resolution (Pick and Buchner, 1979). We emphasise, though, that in our model $\Delta\theta$ is proportionate to $\Delta\phi$ (Snyder, 1979). Therefore, the value of $\Delta\phi$ also sets the spatial resolution according to the latter definition.

Laughlin (1981) summarises four common measures of temporal resolution: frequency response, impulse response, the flicker fusion frequency, and the integration time. In our model, the temporal resolution is represented by the last measure – the integration time Δt of the input neurons. High values of the integration time result in slow responses and, therefore, low temporal resolution. In contrast, short integration time yields a high temporal resolution.

4.2.5 Section summary

In this section we introduced a situated model of target tracking. The model is based on a negative-feedback control system and is implemented in a mobile robot simulator. The input of the model is a distribution of light intensity projected by the target onto the visual sensor, *i.e.*, the model is situated. We showed how the input is processed and transformed into a turning signal of the system using an artificial neural network.

Now that we have introduced the model of target tracking and have identified the representations of spatial and temporal resolution, we introduce an algorithm that maximises the behavioural success of our model.

4.3 Optimisation with an evolutionary algorithm

According to our research method (section 3.3) we aim at maximising the behavioural success of our model by tuning the spatial and temporal resolution to values that

optimise the four determinants selected in section 4.1. To tune the resolution values, an evolutionary algorithm (EA) is employed (Holland, 1975; Goldberg, 1989; Bäck, 1996). We choose an evolutionary algorithm for its ability to find the global maxima of possibly rugged search spaces. Moreover, the algorithm is intuitively clear and requires no extensive knowledge of optimisation techniques for its use.

An evolutionary algorithm employs the principles of natural evolution (*i.e.*, selection and recombination) to optimise a population of candidate solutions for a certain problem. In a standard evolutionary algorithm, each candidate solution is encoded as a string (genotype) containing the parameter values of the solution. In our evolutionary algorithm spatial and temporal resolution ($\Delta\phi$, Δt) are encoded directly as floating-point numbers.

The algorithm is configured as follows. We use rank-based selection. The parents chosen this way generate children by copying (90% of the population) or by cross-over (10% of the population); 80% of the children are then mutated (10% mutation range, per gene). Each population consists of 50 individuals that are initialised to random combinations of spatial and temporal resolution. In each generation, the EA generates 50 children and assesses their fitness. Then, the 75% worst individuals of the population are replaced by the 75% best children. A constant population size is thus maintained. We employ an EA with a highly elitist replacement scheme to accelerate the optimisation process.

An evolutionary algorithm needs a fitness function to calculate the chance of reproduction for the individuals in the population. The fitness function is determined in the following subsection.

4.3.1 The fitness function

In order to assess the fitness of an individual, we determine its behavioural success. For this, we set the values of spatial and temporal resolution in the model of target tracking (the simulated robot) to the values found in the individual's genotype ($\Delta\phi$, Δt). Subsequently, the behaviour of the robot engaged in target tracking is measured as follows. To simulate the target tracking task, a target light source orbits the simulated robot at 100 mm distance (robot's radius = 27.5 mm). For calculation of the behavioural success, we measure the tracking error θ_c (figure 4.2) at every simulation step. The tracking error is averaged over 5 fitness trials that each consist of $100 \cdot \Delta t$ simulation steps. The average error $\bar{\theta}_c$ typically has a standard deviation that falls within 10% of the mean. Finally, the individual's fitness value F is calculated as $F = \theta_{\max} - \bar{\theta}_c$, with $\theta_{\max} = 2$.

In the following section we use the EA to optimise the resolution of the model of target tracking guided by light sensitivity and motion sensitivity.

4.4 Evolution of a trade-off by light sensitivity and motion sensitivity

In this section, we apply the evolutionary algorithm to tune the resolution of the model of target tracking to values that optimise the determinants light sensitivity and motion sensitivity. Our results reveal whether the two determinants cause a trade-off between the spatial and temporal resolution of our model.

In the following subsections, we show how the light sensitivity and motion sensitivity are implemented in the model of target tracking (4.4.1). Subsequently, we outline the details of the experiment in subsection 4.4.2, predict the general shape of the outcomes in subsection 4.4.3, and present the results obtained by applying the evolutionary algorithm in subsection 4.4.4. The results are discussed shortly in subsection 4.4.5, followed by a section summary in subsection 4.4.6.

4.4.1 Implementation in the model

In this section we discuss the implementation of light sensitivity and motion sensitivity in the model of target tracking.

Light sensitivity As was shown in subsection 4.2.4 the activation I of the input neurons depends on the noisy photoreceptor response R' and the integration time Δt (equation (4.3)). Using this fact, the light sensitivity of the system is implemented as follows. After one Δt period, the accumulated activation in each input neuron is compared to an activation threshold ϑ_R (cf. Missler and Kamangar, 1995). When the accumulated activation exceeds the threshold, the neuron is able to engage in the winner-take-all competition (subsection 4.2.4). Therefore, if none of the input neurons' accumulated activations equals or surpasses the threshold, the target is not detected.

Using fixed values of target illumination I_T , response threshold ϑ_R , and angle of incidence $\alpha = 0$ (equation (4.1)), the target is not detected in the following two circumstances. First, if R' is too low, because of a too narrow interommatidial angle $\Delta\phi$. Second, if the integration time Δt is too short. Therefore, the visuo-motor system's light sensitivity is determined by $\Delta\phi$ and Δt (cf. equation (3.2)). The activation threshold is hand-set to an appropriate value such that for every value of I_T used, there exist values for $\Delta\phi$ and Δt that enable at least one input neuron to engage in the winner-take-all competition.

Motion sensitivity The motion sensitivity is implemented in the dynamics of the model of target tracking as is explained here. In subsection 3.1.3 it was explained that the motion sensitivity of biological position-detection systems depends on both the spatial and temporal resolution: the motion sensitivity improves with increasing resolution. In subsection 4.2.4 we showed that, in our model, the spatial and temporal resolution are implemented as the interommatidial angle $\Delta\phi$ and the integration

time Δt , respectively. Therefore, the motion sensitivity depends on the values of $\Delta\phi$ and Δt .

In the following subsection we make four analytical predictions of the optimal resolution values which are used in analysing the results of the evolutionary experiments.

4.4.2 Experimental details

In section 3.2 we showed that the optimal light sensitivity and motion sensitivity depend on illumination and object motion, respectively. Therefore, in the evolutionary experiment we create a number of populations that are assigned different values of the parameters target illumination I_T and angular velocity ω_T . Within a population the parameters have equal values for all individuals. The target's illumination directly sets the required light sensitivity of the system and thus determines the upper bound to spatial and temporal resolution. In contrast, the angular velocity sets the required motion sensitivity and thus creates a lower bound to the product of the resolution values. We use four different values for the target's illumination ($I_T \in \{8, 10, 14, 24\}$ (units)) which were arbitrarily chosen and seven values for the angular velocity ($\omega_T \in \{0.5, 0.7, 1, 2, 4, 6, 8\}$ (10^{-3} rad per simulation step)). Letting 1 simulation step represent 1 ms in real time, the values of ω_T were chosen to represent target velocities that are commonly found in nature (O'Carroll *et al.*, 1996). Altogether we evolve 28 populations.

4.4.3 Analytical predictions

We are able to make four analytical predictions of the evolved values of $\Delta\phi$ and Δt given the values of the control parameters ϑ_R , ω_T , and I_T . The predictions assist in calculating the exact optimal values through situated simulation.

We expect the evolutionary algorithm to optimise the photoreceptor response up to the activation threshold ϑ_R , *i.e.*, where the model is just able to perceive its target. At that threshold, both the light sensitivity and motion sensitivity are expected to be optimal. When optimised, the best-responding neuron's activation I^* , summed over one Δt period, is described as follows.

$$I^* = \sum_{t=t_0}^{t_0+\Delta t} T_{\text{step}} \cdot R'(t) = \vartheta_R \quad (4.7)$$

In equation (4.7) the noisy response R' averages to R when taking means over many time intervals Δt . Substituting R' by equation (4.1) results in the following relation between the resolution values and the control parameters ϑ_R , ω_T , and I_T .

$$T_{\text{step}} \cdot \Delta\phi \cdot \sum_{-\frac{1}{2}\Delta t}^{\frac{1}{2}\Delta t} e^{-2.77 \cdot \left(\frac{\omega_T \cdot t}{c \cdot \Delta\phi}\right)^2} = \frac{\vartheta_R}{I_T} \quad (4.8)$$

with $\omega_T \cdot t = \alpha$ as the angle of incident light, $c \cdot \Delta\phi = \Delta\varrho$ as the photoreceptor acceptance angle, and ϑ_R, I_T as defined before.

Equation (4.8) can not be solved analytically, which yields one of the reasons for modelling the system. Nevertheless, the equation allows us to make four predictions on the experimental results. First, for fixed values of ϑ_R , ω_T , and I_T , there exists a range of values of $\Delta\phi$ and Δt that satisfy equation (4.8). Second, with ω_T fixed, increasing the value of I_T forces the system to decrease the value of at least one of the resolution parameters (and *vice versa*). Third, with I_T fixed, increasing the value of ω_T forces the system to increase the value of at least one of the resolution parameters (and *vice versa*). Fourth, equation (4.8) predicts a trade-off between spatial and temporal resolution.

The experimental results yield the exact influence of I_T and ω_T on the evolved values of $\Delta\phi$ and Δt in the following subsection.

4.4.4 Results

The development of the fitness during the evolutionary optimisation process is shown in figure 4.5 for two populations ($\omega_T = 0.008$, $I_T = 14$ and $\omega_T = 0.001$, $I_T = 14$, respectively; $\vartheta_R = 35$). The results indicate that all 28 populations converge to a stable point in the $(\Delta\phi, \Delta t)$ space within 8 generations (theoretically maximal fitness value equals 2; see subsection 4.3.1). Apparently, the fitness landscape is smooth and an optimal solution is easily found. In figure (a), the fitness reaches lower values than in figure (b). One of the reasons is that the former population is assigned a higher value for ω_T than the latter, which makes smooth pursuit more difficult. Another reason will be given in the section summary (subsection 4.4.6) after the analysis of the results which is performed in the following paragraphs.

The role of the target's angular velocity Since there are only two evolvable parameters ($\Delta\phi, \Delta t$), their values can be plotted at various evolutionary stages in the experiment. In figure 4.6 the mean values of $\Delta\phi$ and Δt of seven populations are depicted at four stages in the evolution (generations 0, 1, 2, and 14). In all populations $I_T = 8$, whereas ω_T varies from 0.0005 to 0.008. In the figures, it is shown that the marks representing the mean resolution values slowly progress to lie on a curve. In figure 4.6 (d), an imaginary curve connecting the marks resembles the general shape of the trade-off curve drawn in section 3.2 (figure 3.2). Therefore, the results suggest that the optimal resolution values corresponding to one value of target light intensity lie on a trade-off curve. The analytically predicted optimal resolution values are described by equation (4.8). The equation, however does not readily describe the shape of the trade-off curves that connect the optimal resolution values. Nevertheless, we know that populations that are assigned equal values for I_T need equal light sensitivity to reach ϑ_R , *i.e.*, they lie on an iso-sensitivity curve in the $(\Delta\phi, \Delta t)$ space (equation (3.3)). Therefore, we fit the datapoints in figure 4.6 (d) to the parameterised inverse square-root function defined by the equation using least-squares minimisation. The fitted curve is shown in figure 4.7 (a) (parameter

(a) $\omega_T = 0.008$, $I_T = 14$ (b) $\omega_T = 0.001$, $I_T = 14$

Figure 4.5: Development of the fitness of two populations during evolution. The values of target illumination and angular velocity are $\omega_T = 0.008$, $I_T = 14$ (a) and $\omega_T = 0.001$, $I_T = 14$ (b).

$k = 1.010 \pm 0.003$, averaged over 5 experiments). The relative fitting error equals 1%. Equation (3.3) is a strictly monotonic decreasing function. Hence, the fitted curve indicates a trade-off between spatial and temporal resolution.

The role of the target velocity ω_T in the evolved trade-off is determined as follows. Figure 4.8 depicts the optimised value of $\Delta\phi$ (a) and Δt (b) as a function of ω_T , averaged over 5 experiments ($I_T = 8$). The figures show that $\Delta\phi$ increases with increasing angular velocity ω_T , whereas Δt decreases. Therefore, spatial resolution is traded for temporal resolution when the angular velocity is raised.

In summary, we find that, for constant target illumination I_T , the optimal values of $\Delta\phi$ and Δt lie on a trade-off curve. The target velocity ω_T determines the trade-off point, *i.e.*, the location on the trade-off curve, to which a population evolves.

The role of the target's illumination When varying the target illumination I_T , we obtain the results shown in figure 4.7 (b). In the figure we present the evolved mean values of spatial and temporal resolution in the 28 populations. An inverse square-root function (equation (3.3)) is fitted to the datapoints corresponding to each value of I_T used. As a result we obtain scaled versions of the trade-off curve. The values of the fitted scaling parameter k are plotted in figure 4.9 as a function of illumination I . The relative fitting error (averaged over 5 experiments) does not exceed 1% for any of the curve fittings. In the figure, it is shown that the value of k , and thus the values of $\Delta\phi$ and Δt , decrease when the illumination is increased. Therefore, the optimal resolution is higher when the target is well illuminated than when the target is dimly lit. In all curves shown in figure 4.7 (b), we observe the

(a) Generation 0

(b) Generation 1

(c) Generation 2

(d) Generation 14

Figure 4.6: Mean values of $\Delta\phi$ and Δt in 7 populations at various stages in the experiment. In all populations $I_T = 8$, whereas ω_T varies from 0.0005 to 0.008 (see legend).

same velocity dependency: for all values of I_T measured, ω_T determines the trade-off point. In summary, we find that the target illumination, I_T , determines the position of the trade-off curve, *i.e.*, the scaling factor k (subsection 3.2.2).

4.4.5 Discussion

When comparing the predicted results to the experimental results, it is observed that the predictions are a generalised description of the results. In fact, we find that for fixed values of ω_T and I_T , the resolution parameters do not converge to a range of values (as was predicted in subsection 4.4.3), but to a trade-off point instead. Analysis of the evolved behaviour reveals the following reason for this

(a)

(b)

Figure 4.7: (a) The trade-off curve for $I_T = 8$. (b) The trade-off curves for I_T varying from 8 to 24 (see legend). ω_T varied from 0.0005 to 0.008 as in (a).

(a)

(b)

Figure 4.8: The optimised values of $\Delta\phi$ (a) and Δt (b) as function of target velocity ω_T . In both figures, $I_T = 8$. All datapoints were averaged over 5 experiments.

observation. The model of target tracking seems to adapt its spatial and temporal resolution in such a way that during one integration time the target will move well within the confines of a single receptor's receptance angle. For instance, with $I_T = 8$ and $\omega_T = 0.004$ rad per simulation step, the system's visual resolution converged to $\Delta\phi = 0.19$ rad, $\Delta t = 29$ steps ($\Delta\varrho = \Delta\phi = 0.19$ rad). The angular distance travelled by the target during one time interval Δt equals 0.12 rad. However, decreasing Δt causes $\Delta\phi$ to grow (see equation (4.8)) and large values of $\Delta\phi$ result in a large target localisation error. Therefore, Δt does not shrink infinitely and the system converges to a trade-off point.

Figure 4.9: The fitted scaling parameter k (equation (3.3)) as a function of the illumination I .

Moreover, from the results it appears that when increasing ω_T , Δt is lowered, while $\Delta\phi$ grows larger. For all ω_T , equation (4.8) is satisfied. As mentioned before, lowering Δt enhances response speed, but increasing $\Delta\phi$ enlarges the target localisation error. This is reflected in the maximal fitness obtained in a different population for different values of ω_T . In figure 4.5, the highest angular velocity results in the lowest maximal fitness due to increased $\Delta\phi$.

Finally, changes in target illumination I_T result in simultaneous increase or decrease of the resolution parameters. We note again that these observations could not be derived from analytic studies directly (see equation (4.8)), but only follow from the simulations.

4.4.6 Section conclusion

In summary, to maximise the behavioural success of our model of target tracking, we tuned the spatial and temporal resolution to values that optimise the determinants light sensitivity and motion sensitivity. We found that these values are represented by points on a trade-off curve. The scale of the curve and the location of the point on the curve are determined by the illumination and the angular velocity of the target. Therefore, we may conclude that the two determinants cause a trade-off between the spatial and temporal resolution of our model of target tracking.

4.5 Evolution of a trade-off by heat production and motion sensitivity

In this section we investigate whether the determinants heat production and motion sensitivity cause a similar trade-off as the determinants investigated in the previous section.

In subsection 3.2.5 it was argued that the capacity of the heat-dissipation system sets the optimal level of heat production and thus of the visual resolution. Therefore, to show how the determinant heat production causes a trade-off, we vary the capacity of the model's heat-dissipation system. For each capacity value used, we let the model's resolution evolve to the optimal values. The evolved values are subsequently fitted to a trade-off curve, comparably to the procedure of section 4.4.

The section is organised in a similar fashion as section 4.4. In subsection 4.5.1, we explain how the determinant heat production is implemented in the model. Thereafter, we summarise the experimental details in subsection 4.5.2; make predictions of the outcomes of the experiments in subsection 4.5.3; and present the results in subsection 4.5.4. Finally, we summarise the section in subsection 4.5.5.

4.5.1 Implementation in the model

The visual system in the model of target tracking is augmented by adding a parameter *heat counter* (hc) that keeps track of the temperature in the neural system, caused by metabolic activity. As was argued in subsection 3.2.5, heat is a critical factor in the functioning of the brain (Baker, 1979; Cabanac and Brinnet, 1985; Falk and Gage, 1998) – overheating quickly leads to malfunction or brain death. In accordance with this observation we define a threshold ϑ_H to indicate the maximal heat permitted in the visuo-motor system. Once the value of hc exceeds the threshold ϑ_H the system stops performing target-tracking behaviour which results in a low fitness value in the evolutionary optimisation process.

The heat of the system is updated as follows. We first recall that in the model, the activation of the input neurons is flushed after each time interval Δt to represent the operation of the ion pumps (subsection 4.2.4). To model the heat production in the artificial visual system, we increment the parameter hc after each time interval Δt by a value proportional to the number of photoreceptors.

$$hc(t + T_{\text{step}}) = hc(t) + H \cdot N_{\text{sensors}} \quad \text{after each time interval } \Delta t \quad (4.9)$$

with T_{step} indicating the duration of one simulation step. The constant H represents the heat produced by a photoreceptor due to the metabolic activity of its ion pumps. To model heat dissipation (*i.e.*, brain cooling) a variable heat decay d_H ($0 < d_H < 1$) reduces the value of hc at every simulation step. We model the heat-decay process as an exponential function $hc(t) = hc_0 \cdot e^{-a \cdot t}$ (see *e.g.*, Edwards and Penney, 1982). In the experiments we use the discretized, recursive, version of this function expressed

in equation (4.10).

$$hc(t + T_{\text{step}}) = (1 - d_H) \cdot hc(t) \quad \text{at each simulation step} \quad (4.10)$$

with $d_H = 1 - e^{-a \cdot T_{\text{step}}}$ and $hc(0) = hc_0$.

4.5.2 Experimental details

We assign five different values to the heat decay factor d_H , linearly spaced between 0.002 and 0.006. The values were chosen in such a way to match the time scale. Each value is assigned to a set of populations in the evolutionary algorithm, in a similar way in which we assigned values of the target's light intensity to the populations in the previous experiment (see section 4.4). The target's velocity takes the same seven values as before ($\omega_T \in \{0.5, 0.7, 1, 2, 4, 6, 8\}$ (10^{-3} rad per simulation step)) thus resulting in $5 \times 7 = 35$ populations. To eliminate the role of light sensitivity in the trade-off, the light intensity of the target is set to a high value.

The same parameters for the evolutionary algorithm (*e.g.*, mutation rate, population size) were used as in the previous experiments on light sensitivity (section 4.4).

4.5.3 Analytical predictions

We predict the position of the trade-off curve, from the heat decay factor d_H . When optimised, the system is expected to reach a steady state as depicted in figure 4.10.

Figure 4.10: The temperature in the model of target tracking over time. After every interval of Δt simulation steps the temperature increases due to heat production in the neurons (see equation (4.9)). After every single simulation step the temperature decays exponentially (see equation (4.10)).

The optimal steady state is mathematically described as follows.

$$\vartheta_H \cdot (1 - d_H)^{\Delta t} + a = \vartheta_H \quad (4.11)$$

with ϑ_H (heat threshold) and d_H (heat-decay factor) as described above and a denoting the amount of heat added after each integration time Δt : $a = N_{\text{sensors}} \cdot H = (\frac{\pi}{\Delta\phi} + 1) \cdot H$ (see also equation (4.6)). Below, equation (4.11) is rewritten to express the relation between the optimal values of $\Delta\phi$ and Δt .

$$\Delta\phi = \frac{\pi}{\frac{\vartheta_H}{H}(1 - (1 - d_H)^{\Delta t}) - 1} \quad (4.12)$$

Equation (4.12) is a monotonically decreasing function and, therefore, describes trade-off curves in the $\Delta t, \Delta\phi$ plane. The parameters ϑ_H and H are kept constant in all populations in the evolutionary algorithm ($\vartheta_H = 300$, $H = 2$, chosen to match the values of d_H). Hence, the position of the curves only depends on d_H .

4.5.4 Results

Comparably to the results of section 4.4, the fitness landscape appears smooth and the optimal values for Δt and $\Delta\phi$ are easily found within 8 generations. The average values of spatial and temporal resolution evolved in each population are shown in figures 4.11 (a) ($d_H = 0.003$, different values for ω_T) and (b) (different values for d_H and ω_T).

(a)

(b)

Figure 4.11: The trade-off curves induced by heat production.

The figures also show the fitted trade-off curves described by equation (4.12). In order to measure how well the evolved data are predicted by equation (4.12) we calculate the fitting error. The relative fitting error averaged over all 35 populations and over the six repetitions of the experiment equals 5%. The fitting error suggests that the differences between the evolved values and the values predicted by equation (4.12) are not significant.

4.5.5 Section conclusion

In summary, the values of spatial and temporal resolution that optimise the determinants heat production and motion sensitivity are represented by points on a trade-off curve.

As in section 4.4.4 we find that the angular velocity of the target determines the position of the trade-off point to which a population has evolved (see figure 4.11 (a)). We stress that this finding cannot be predicted from equation (4.12). Moreover, from figure 4.11 (b) we find that an increase of the heat-decay factor scales the trade-off curves toward a higher spatial and temporal resolution as can be predicted through equation (4.12). The heat-decay factor d_H fulfils a similar role in the trade-off as light intensity I_T (see section 4.4.4). In conclusion, with sufficient illumination of the target, *i.e.*, when light sensitivity is not a limiting factor, the determinants heat production and motion sensitivity as well cause a trade-off between spatial and temporal resolution in the model of target tracking.

4.6 Evolution of a trade-off by energy consumption and motion sensitivity

Energy consumption was the last determinant identified in section 3.2. In this section it is verified whether the determinants energy consumption and motion sensitivity cause a trade-off between spatial and temporal resolution in the model of target tracking. In analogy to section 4.5, we show that the position of the trade-off curves is predicted by a mathematical equation that describes the process of energy supply and energy consumption. Since the results of the current evolutionary experiments are similar to those in section 4.5, we restrict ourselves to providing the implementation of the determinant energy consumption in the model (subsection 4.6.1) and the analytical predictions (subsection 4.6.2).

4.6.1 Implementation in the model

The determinant energy consumption is implemented in the model of target tracking as follows. We assume a mean rate of energy input to the retina e_{input} (energy units per simulation step). The rate depends on the capacity of the blood vessels leading to the eye and on the amount of energy available in the body. As stated in subsection 3.2.6, in rest, the rate can be as high as 10% of the energy available in the

blowfly (Laughlin *et al.*, 2000), or even 20% in humans (Kety, 1957). Furthermore, in analogy to section 4.5, we define a constant e_{ionpump} as the energy consumed by a photoreceptor due to the metabolic activity of its ion pumps during a single time interval Δt .

4.6.2 Analytical predictions

In the optimal steady-state circumstances, the inflow of energy per integration time matches the energy consumed by all photoreceptors as expressed in equation (4.13).

$$e_{\text{input}} \cdot \Delta t = e_{\text{ionpump}} \cdot N_{\text{sensors}} \quad (4.13)$$

After substituting the number of photoreceptors N_{sensors} by equation (4.6) we arrive at the following relation between spatial and temporal resolution.

$$\Delta\phi = \frac{\pi}{\frac{e_{\text{input}} \cdot \Delta t}{e_{\text{ionpump}}} - 1} \quad (4.14)$$

With $e_{\text{input}} > e_{\text{ionpump}}$, equation (4.14) is monotonically decreasing when Δt is larger than one simulation step. Hence, equation (4.14) describes a set of trade-off curves, scaled by the fraction $\frac{e_{\text{input}}}{e_{\text{ionpump}}}$. In analogy with the results in subsections 4.4.4 and 4.5.4 we expect that the spatial and temporal resolution in the model of target tracking evolves toward a trade-off curve described by equation (4.14). Moreover, the angular velocity of the target is expected to determine the evolved position of the trade-off point.

4.6.3 Section conclusion

Based on the analysis above, we expect that the resolution values that optimise the energy consumption and motion sensitivity are represented by a point on a trade-off curve. The position of the point on the curve is determined by the angular velocity of the target, whereas the scale of the trade-off curve is set by the fraction $\frac{e_{\text{input}}}{e_{\text{ionpump}}}$. Based on this expectation, we conclude that the determinants energy consumption and motion sensitivity cause a trade-off between spatial and temporal resolution in our model of target tracking.

4.7 Comparison to biological systems

Below, we compare our results to biological visual systems. In sections 4.4, 4.5, and 4.6 we verified that the determinants light sensitivity, motion sensitivity, heat production, and energy consumption all cause a trade-off between the spatial and temporal resolution of our model of target tracking. The optimal resolution values (*i.e.*, those that maximise the behavioural success of our model) are determined by four features of the visual ecology: (1) the target's illumination; (2) the target's

velocity; (3) the heat dissipation; and (4) the energy supply. In comparison, we determine the role of these features in setting the resolution values of biological visual systems (in subsections 4.7.1, to 4.7.4, respectively). A summary follows in subsection 4.7.5.

4.7.1 The role of the target's angular velocity

O'Carroll *et al.* (1996) compared the target-tracking behaviour of bumblebees and humans. They showed that bumblebees are optimised for higher image velocities (350 deg/s) than human are (2 deg/s). This fact mirrors the visual ecology of the animals: in their everyday environment, the fast-flying bumblebees are typically confronted with higher values for ω_T than humans. Presumably, during evolution bumblebees developed a high motion sensitivity since their low spatial resolution allowed them to evolve high temporal resolution. A similar argument applies to the fly *Fannia canicularis* (Land and Collett, 1974) where the low spatial resolution of compound eyes allows for high temporal resolution without compromising the light sensitivity. Finally, Laughlin and Weckström (1993) found that in a set of Dipteran species (true flies) "the speed of photoresponse [*i.e.*, temporal resolution] is associated with the speed of image motion encountered on the retina". The examples presented here provide evidence that, in congruence with our results, the target's velocity sets the position of an animal's resolution on a trade-off curve.

In addition, we show that the differences in visual resolution in fovea and the peripheral parts of the retina (subsection 1.2.2) may also be caused by the determinant motion sensitivity. In humans, as well as in certain fly species, the relatively high spatial resolution in the fovea is employed for smooth pursuit of slowly moving objects (low ω_T). In contrast, the highly motion-sensitive periphery primarily serves the detection of fast moving objects (high ω_T) (Young, 1971; Collett and Land, 1975). Again, we find that the target's velocity sets the trade-off point in biological systems.

In conclusion, the examples suggest that the target's angular velocity determines the position of the trade-off point not only in our model, but also in biological target tracking. The motion sensitivity of visual systems is one of the determinants that may cause a trade-off in biological systems.

4.7.2 The role of the target's illumination

In this subsection we show that in many biological visuo-motor systems the target's illumination scales the trade-off curve on which the systems' resolution values lie. The scaling is accomplished by means of dark-adaptation mechanisms of which we discuss two examples (see Pick and Buchner, 1979; Snyder, 1979; Warrant, 1999). First, neural pooling (clustering the neural output of multiple ommatidia) is a natural strategy to optimise the light sensitivity. This strategy reduces the spatial resolution without increasing (or decreasing) the temporal resolution of the animal, *i.e.*, it scales the trade-off curve. Second, many animals adopt temporal summa-

tion, *i.e.*, elongating Δt , as a dark adaptation strategy. In both cases, the target illumination scales the optimal resolution of biological systems comparably to the artificial model.

Moreover, when comparing the visual ecology of hoverflies, that live in brightly illuminated environments, to the ecology of daddy-longlegs (dimly lit environments) we find another similarity with the model of target tracking. Both the spatial and temporal resolution in hoverflies are higher than in daddy-longlegs (Laughlin, 1996). Therefore the trade-off curve corresponding to animals living in brightly lit environments is scaled with respect to animals living in dark environments. This relation was predicted by our studies using the model of target tracking.

We conclude that the target's illumination scales the trade-off curve of biological systems, in analogy to our experiments. Consequently, our results predict that light sensitivity causes a trade-off in biological systems, as it does in the model of target tracking.

4.7.3 The role of heat dissipation

The third comparison concerns the heat dissipation of the model of target tracking and of biological visuo-motor systems. In subsection 4.5.1 we modelled the capacity of the heat-dissipation system in terms of the heat decay (d_H).

The heat-decay factor in the model is functionally equivalent to the capacity of the heat-dissipation system in an animal's brain: given a fixed level of metabolic activity, both the heat-decay factor (in the model) and the capacity of the heat-dissipation system (in the brain) determine the steady-state temperature of the visuo-motor system (see subsection 3.2.5 and Baker, 1979; Cabanac and Brinnet, 1985). Our results show that in the model of target tracking, the heat decay scales the trade-off curve (subsection 4.5.4). In congruence, the capacity of an animal's heat-dissipation system limits spatial and temporal resolution in the visual system (see subsection 3.2.5). Hence, our studies predict that the capacity of the heat-dissipation system scales the trade-off curve in biological visuo-motor systems: more dissipation allows higher heat production and thus improved spatial and temporal resolution as is reflected in the position of the trade-off curve.

4.7.4 The role of energy supply

Finally, we compare the energy supply of the model and of biological systems. In subsection 4.6.1 we modelled the energy supply of the model as the variable e_{input} .

The energy supply to the model's visual system was shown to yield scaled versions of a trade-off curve. Laughlin *et al.* (2000) provide a set of four fly species in which a similar phenomenon can be observed. The mass of the eye, relative to the body mass, is different in daddy-longlegs (smallest), blowflies (intermediate), and hoverflies (largest). Furthermore, it is argued that the mass of a neural system is a simple measure of the system's metabolic activity (heat production and energy consumption). Moreover, in hoverflies, spatial and temporal resolution is higher

than in blowflies (cf. Land and Collett, 1974 and Collett and Land, 1975) which in turn exceeds the resolution in daddy-longlegs (Laughlin, 1996). Therefore, in these biological visuo-motor systems, too, the resolution is scaled by the energy supply to the retina.

4.7.5 Section summary

In summary, we compared the roles of four features of the visual ecology (the target's illumination; the target's velocity; the heat dissipation; and the energy supply) in setting the resolution of our model and of biological visual systems. From the comparison we may conclude that the four determinants investigated above cause a trade-off in our model, as well as in animals.

4.8 Chapter discussion and conclusions

In this section we discuss the results of the experiments with our model of target tracking. First, we determine to what extent our results can be interpreted as the cause of the trade-off in biological systems in subsection 4.8.1. Moreover, in subsection 4.8.2, we point at the limitations and possible improvements of the model of target tracking. Finally, chapter conclusions are drawn in subsection 4.8.3.

4.8.1 Application to biological systems

To apply our results to biological systems we determine whether our findings are specific to the peculiarities of our model of target tracking. In our model we implemented a 1-dimensional eye (subsection 4.2.4), whereas the equations describing the receptor response (equation (4.1)) of biological eyes are based on a 2-dimensional retina. Therefore, our results cannot be directly compared to biological data. However, the qualitative pattern of our results, *i.e.*, the trade-off curves and the dependency on the determinants investigated, is reflected in the resolution of biological visual systems (see section 4.7). Moreover, the visual system in the model of target tracking was designed to model biological visual systems (see subsection 4.2.4) and is subject to the same physical limiting factors as animals (*i.e.*, light sensitivity, motion sensitivity, heat production, and energy consumption; see also section 3.2). Therefore, we expect that our results apply to biological systems (see also chapter 7).

4.8.2 Limitations and possible improvements of the model

The model of target tracking is limited in several respects. In particular, the neural model underlying target-tracking behaviour is too simple from a biological perspective since it is constrained to a model of the retina. Moreover, besides the use of position-detection mechanisms, many animals employ motion-detection systems to perform target tracking (see subsection 4.2.1). In fact, motion detection is central to the perception of stable and moving objects in insects and humans. In the fly, motion

detectors govern the target-tracking and course-stabilisation behaviour (*e.g.*, Poggio and Reichardt, 1976; Egelhaaf and Borst, 1993a). Also in human vision, motion-detector cells play an important role, such as when differentiating between motion induced by eye, head, and body movements and between motion and non-motion of objects (*e.g.*, Murakami and Cavanagh, 1998).

Besides limitations in the simulated neural system, we used a simplified motor model as well. For instance, we did neither model the momentum of the system, nor was it implemented in the robot simulator. Moreover, the system latency (20–40 ms in the target-tracking behaviour of flies; Wehrhahn, Poggio, and Bülthoff, 1982) was not modelled. Both factors, momentum and latency, play an important role in determining the stability of the control system upon which the model is based (figure 4.1). As will be shown in chapter 5, the resolution values might optimise the stability as well.

Moreover, the specific way of implementing light sensitivity, heat production, and energy consumption in the system resembles the actual processes in natural systems only at a very abstract level.

These limitations constrain the biological plausibility of our artificial visuo-motor system. They are caused by the fact that our study compares different animal species that use different neural mechanisms and different locomotory systems. Therefore, we suggest to focus further research on finding general principles in the design of biological visuo-motor systems, upon which we base the design of a generic artificial visuo-motor system. These principles include photoreceptor response (Laughlin, 1981), small-field and wide-field motion detection (Egelhaaf and Borst, 1993a; Pack *et al.*, 2001), spatial contrast sensitivity (Thibos, 2000) and temporal contrast sensitivity (Geissler and Albrecht, 2000).

4.8.3 Chapter conclusions

Our research has led us to the following two conclusions. First, by evaluating our results we conclude that four of the five determinants (light sensitivity, motion sensitivity, heat production, and energy consumption) cause a trade-off between the spatial and temporal resolution in artificial position-detection systems. In such systems, the four determinants lead to optimal resolution values that lie on a trade-off curve when maximising their behavioural success. The fifth determinant (contrast sensitivity) was not investigated but is expected to yield the same results. Second, by analogy we tentatively conclude that the four determinants cause scaled versions of trade-off curves in the evolved resolution values of biological position-detection systems.

Chapter 5

The resolution of motion-detection systems

“Never mistake motion for action.”

Ernest Hemingway (1899 – 1961), U.S. novelist.

Chapter contents: resolution optimisation, motion-detection systems, gaze-stabilisation behaviour, visual ecology.

Motion detection is the second principal control system of visually guided behaviour (subsection 3.1.3 and see Zanker and Zeil, 2001b; Euler, Detwiler, and Denk, 2002). In this chapter we study the resolution of motion-detection systems.

5.1 Aim of the chapter

We aim to determine the causes of a trade-off between the spatial and temporal resolution of motion-detection systems. In section 3.2 five determinants were identified, of which we investigate three in the context of motion-detection systems: motion sensitivity, light sensitivity, and contrast sensitivity. The remaining two determinants (*i.e.*, heat production and energy consumption) are not investigated for the following reason. The two determinants were studied in the context of position-detection systems in chapter 4. Whether in position-detection systems or in motion-detection systems, high spatial resolution leads to large numbers of neurons that produce much heat and consume much energy. In analogy, high temporal resolution leads to high heat production and energy consumption, irrespective of the type of detection mechanism. Therefore, we expect that the role of heat production and energy consumption in position-detection systems can be generalised to motion-detection

systems. In contrast, we will show that the light sensitivity and motion sensitivity of position-detection systems are quite different from motion-detection systems, which makes the determinants interesting for another investigation here.

As a second aim of the chapter, we improve the biological plausibility of our model (subsection 4.8.2) by studying contrast sensitivity as well. Comparably to biological systems we will exploit the statistics of natural visual scenes in our model (cf. Dror, O'Carroll, and Laughlin, 2001). Because of this exploitation, we may gain a better understanding of the trade-off as a natural phenomenon. Natural scenes are characterised by different factors, such as the mean illumination and average contrast. Therefore, it is interesting to study, besides light sensitivity, the contrast sensitivity of motion-detection systems.

5.1.1 Method

As in the previous chapter, we investigate a situated model of visually guided behaviour. Target-pursuit behaviour was the model behaviour in chapter 4. Here, we construct a model of gaze stabilisation, a typical behaviour that relies on motion detection and that is omnipresent in the animal world (subsection 3.1.3).

To determine whether the selected determinants cause a trade-off in the model, we maximise its behavioural success by tuning the spatial and temporal resolution to values that optimise the determinants (see our research method in section 3.3). If the tuned resolution values lie on a trade-off curve, we conclude that the determinants cause the trade-off.

In particular, in this chapter, we tune the optimal resolution values indirectly. For this, we establish mathematical relations that describe how the spatial and temporal resolution set the value of the determinants. Then, we optimise the determinants directly to optimise the behaviour of the model (*i.e.*, to yield a maximal behavioural success). Using the aforementioned relations, we derive the optimal resolution values, and the eventual existence of trade-off curves, from the optimal values of the determinants.

5.1.2 Chapter survey

The structure of the current chapter reflects, to a large extent, the structure of chapter 4. We introduce our model of gaze stabilisation in section 5.2. Thereafter, we perform experiments that verify the causes of the trade-off in four sections. First, we show how the resolution sets the model's motion sensitivity in section 5.3. Second, in section 5.4 we determine the optimal motion sensitivity. Third and fourth, the same procedure is employed to verify whether the determinants light sensitivity and contrast sensitivity cause a trade-off in sections 5.5 and 5.6. Having completed the experiments, we determine the causes of the trade-off between the spatial and temporal resolution of the model of gaze stabilisation in section 5.7. In section 5.8, we discuss the results and summarise the chapter.

5.2 A situated model of gaze stabilisation

In this section, we present a situated model of gaze-stabilisation behaviour based on the theory and experimental data of biological visual systems. We emphasise that we do not aim at providing a new model of gaze-stabilisation behaviour per se; rather, we present our model as one that is biologically plausible and that is ideally suited to study the trade-off.

In subsection 5.2.1, we describe the components of the model. Thereafter, in subsection 5.2.2, we elaborate on the most interesting component: the augmented Reichardt correlator, which is used to detect motion of moving visual scenes. The augmented Reichardt correlator is characterised by two response functions: the frequency-dependent response function and the velocity-dependent response function. The functions are established in subsections 5.2.3 and 5.2.4, respectively. A summary is presented in subsection 5.2.5.

5.2.1 Components of the model

Our model is inspired by a generic model of rotational gaze-stabilisation behaviour (Land, 1992). The behaviour involves corrective shifts of gaze in the horizontal plane in response to wide-field, rotational visual motion. Land's model is a delayed negative feedback system that captures many aspects of the gaze-stabilisation behaviour shown by animals. However, the model is not situated. Moreover, the spatial and temporal resolution are not represented in the model. To overcome these shortcomings, we add three components to Land's model. The additional components are inspired by various other models and theoretical and experimental findings (Heisenberg and Wolf, 1984; Egelhaaf and Borst, 1993b; Warzecha and Egelhaaf, 1996; Huber and Bühlhoff, 1998; Neumann and Bühlhoff, 2001; Sandini, Panerai, and Miles, 2001) and comprise a Reichardt correlator, a low-pass filter, and a source of noise. The resulting control diagram of our model of gaze stabilisation is shown in figure 5.1. The variable V_b represents the rotational angular velocity

Figure 5.1: Control diagram for gaze-stabilisation behaviour.

of the visual background, henceforth called background velocity, while V_s denotes the system's response velocity and $V_r (= V_b - V_s)$ represents the retinal slip velocity observed by the system. Furthermore, a Reichardt correlator detects the velocity V_r and the correlator's output is fed into a temporal low-pass filter (LP) represent-

ing the integrating properties of the system (Heisenberg and Wolf, 1984; Warzecha and Egelhaaf, 1996; Sandini *et al.*, 2001). To simulate the response variance of biological neural systems (Pelli, 1990) we add noise to the LP-filtered signal. The factors *Delay* and *K* represent the system's latency and an internal gain constant, respectively (Land, 1992).

The most complex part of our model is the Reichardt correlator. In subsection 3.1.3 we outlined the operation of a symmetric correlator. To study the trade-off between spatial and temporal resolution, we present an augmented version of the symmetric correlator in the next subsection.

5.2.2 The augmented Reichardt correlator

The symmetric Reichardt correlator model (figure 3.1.3; Reichardt, 1961) is a simplified description of the mechanisms for motion detection in biological visual systems. To represent a more biologically plausible mechanism for motion detection we introduce the augmented Reichardt correlator. Biological motion detectors do not multiply raw intensity values as is suggested by the model in subsection 3.1.3. Instead, the light intensity values are transduced by photoreceptors and transformed by various optical and neural pre-filters (Snyder, 1979; Land, 1981; Snippe and Koenderink, 1994; Dror *et al.*, 2001). We augment the basic Reichardt correlator by simulating photoreception and by employing spatial and temporal band-pass pre-filters, *G* and *H*, respectively (Van Hateren, 1992b). Our augmented Reichardt correlator model is shown in figure 5.2. Below we provide a mathematical description of the model. We start by describing the process of photoreception, then the design of the pre-filters, and finally the correlation process.

Figure 5.2: Augmented symmetric Reichardt correlator. Also shown are the optical and neural pre-filtering stages performed in the early stages of visual processing.

Photoreception

We model photoreception as a linear transformation of light intensity values to a neural receptor response signal. There are two reasons for defining photoreception as a linear transformation. First, the mean number of photons absorbed by a photoreceptor is known to be proportional to the luminance of the visual scene, *i.e.*, light intensity (Snyder, 1979). Second, we assume that, in an adapted state, the receptor response is a linear function of the number of absorbed photons. In the following paragraph, we describe the response of the pre-filters in our model.

Pre-filters

In our model, the input signal is first spatially band-pass filtered (G). In natural systems, spatial band-pass filtering takes place in the optics of the eye (Snyder, 1979) and the dendritic trees of the neural mechanisms processing visual information. We model the spatial pre-filtering by applying a difference of Gaussians (DOG) filter to the input pattern (Rodieck, 1965). The DOG filter is characterised by the half-widths of the two Gaussian functions involved, $\Delta\varrho_1$ and $\Delta\varrho_2$. The amplitude response is calculated as the absolute value of the Fourier transformed DOG filter:

$$|G(f_s)| = \exp\left(-\frac{1}{2}(\Delta\varrho_1)^2 \cdot (2\pi \cdot f_s)^2\right) - \exp\left(-\frac{1}{2}(\Delta\varrho_2)^2 \cdot (2\pi \cdot f_s)^2\right) \quad (5.1)$$

with $\Delta\varrho_1 < \Delta\varrho_2$ and with f_s representing the spatial frequency of the input pattern. The complete derivation of all filter response functions are provided in appendix A. The filter G is symmetric and therefore has zero phase response.

After applying spatial filtering we employ temporal band-pass (BP) filtering (H in figure 5.2). In natural visual systems, the cell membranes and ionic channels of photoreceptors and neurons act as resistor-capacitor networks (first-order filters) that temporally low-pass filter the incoming signals (Softky and Koch, 1995; Laughlin, 1996). High-pass temporal filtering is accomplished in natural systems by neural units of the differentiating type such as the on-off cells in the fly lamina (Arnett, 1972; Dror *et al.*, 2001) and in the vertebrate retina (Fiorentini *et al.*, 1990). The filter output is calculated by consecutively applying a low-pass (H_{LP}) and high-pass (H_{HP}) filter. The filters are implemented as first-order RC-circuits (Oppenheim *et al.*, 1997). The amplitude response $|H|$ of the BP filter is calculated as the product of the amplitude responses of the two individual filters $|H_{LP}|$ and $|H_{HP}|$ with filter constants τ_1 and τ_2 , respectively (equations (5.2) to (5.4), see also appendix A).

$$|H_{LP}(f_l, \tau)| = \frac{1}{\sqrt{1 + (2\pi f_l \tau)^2}} \quad (5.2)$$

$$|H_{HP}(f_l, \tau)| = \frac{2\pi f_l \tau}{\sqrt{1 + (2\pi f_l \tau)^2}} \quad (5.3)$$

$$|H(f_l)| = |H_{LP}(f_l, \tau_1)| \cdot |H_{HP}(f_l, \tau_2)| \quad (5.4)$$

The BP filter's phase response $\angle H$ equals the sum of the phase responses of the individual filters (equation (5.5) to (5.7), see also appendix A).

$$\angle H_{LP}(f_t, \tau) = -\arctan(2\pi f_t \tau) \quad (5.5)$$

$$\angle H_{HP}(f_t, \tau) = \frac{1}{2}\pi - \arctan(2\pi f_t \tau) \quad (5.6)$$

$$\angle H(f_t) = \angle H_{LP}(f_t, \tau_1) + \angle H_{HP}(f_t, \tau_2) \quad (5.7)$$

The photoreceptor and filtering stages are our contributions to the augmented Reichardt correlator, as described above. In the following paragraph we describe how the components of the original correlator are incorporated in our model.

The original correlator

The temporal filters in the original Reichardt correlator, I and D , are also modelled as first-order filters, as is usually done in modelling studies (Egelhaaf and Borst, 1989; Egelhaaf, Borst, and Reichardt, 1989). The filters are expressed as $I(f_t) = H_{LP}(f_t, \tau_3)$ and $D(f_t) = H_{HP}(f_t, \tau_4)$. The variables τ_1 to τ_4 may all have different values.

To model the process of correlation, the pre-filtered left-hand and right-hand input signals are multiplied in the last stage of the Reichardt correlator. The left multiplication unit's output is sensitive to rightward motion, while the right multiplication unit's output is sensitive to leftward motion (see subsection 3.1.3). Subtracting the output signals of the two multiplication units reduces the contribution of time-dependent components to the response. Moreover, the subtraction phase decreases the sensitivity to uniform illumination changes (Egelhaaf *et al.*, 1989).

All components of the augmented Reichardt correlator have now been introduced. Above, we characterised the pre-filters by their amplitude response functions and phase response functions. Below, we characterise the complete correlator by establishing two response functions: the frequency-dependent response function (subsection 5.2.3) and the velocity-dependent response function (subsection 5.2.4).

5.2.3 The frequency-dependent response function

In this subsection we establish the frequency-dependent response function of the augmented Reichardt correlator. The function is used to calculate the resolution of the model of gaze stabilisation and therefore plays an important role in studying the trade-off.

To determine the frequency-dependent response function, we consider a moving sine-wave grating with spatial frequency f_s and angular velocity V . The temporal frequency of the pattern is thus defined as $f_t = V \cdot f_s$. We use I_0 as the mean

luminance level of the grating and ΔI as the grating's modulation amplitude. The spatio-temporal luminance pattern $I(x, t)$ associated with the grating is expressed as:

$$I(x, t) = I_0 + \Delta I \cdot \sin(2\pi \cdot (f_s x - f_t t)) \quad (5.8)$$

with x representing the horizontal position and t taken as the time. We do not include a vertical position since we restrict ourselves to motion in the horizontal plane.

We set the position of the left input arm in figure 5.2 to $x_L = 0$. When using an angular receptor distance $\Delta\phi$, the (angular) position of the right input arm then equals $x_R = \Delta\phi$. The input signals I_1 and I_2 of the left-hand multiplication unit in figure 5.2 are expressed as:

$$I_1(f_s, f_t, t) = |G(f_s)| \cdot |H(f_t)| \cdot |I(f_t)| \cdot I\left(0, t + \frac{\angle H(f_t) + \angle I(f_t)}{2\pi f_t}\right) \quad (5.9)$$

$$I_2(f_s, f_t, t) = |G(f_s)| \cdot |H(f_t)| \cdot |D(f_t)| \cdot I\left(\Delta\phi, t + \frac{\angle H(f_t) + \angle D(f_t)}{2\pi f_t}\right) \quad (5.10)$$

Expressions for the inputs I_3 and I_4 of the right-hand multiplication unit are obtained by substituting I_1 and I_2 by I_3 and I_4 , respectively, and exchanging the filters I and D in equations (5.9) and (5.10).

The time-averaged frequency-dependent response of the correlator is calculated as $R_f = \overline{I_1 \cdot I_2 - I_3 \cdot I_4}$, with the bar indicating space and time averaging (Snippe and Koenderink, 1994). Using elementary trigonometrical rules the response R_f is calculated as (cf. Egelhaaf *et al.*, 1989; Zanker, Srinivasan, and Egelhaaf, 1999):

$$R_f(f_s, f_t, \Delta I) = P(f_s, f_t) \cdot \Delta I^2 \cdot \sin(2\pi \cdot f_s \cdot \Delta\phi) \cdot \sin(\angle D(f_t) - \angle I(f_t)) \quad (5.11)$$

In the formula, P represents the amplitude response of the spatio-temporal filters:

$$P(f_s, f_t) = |G(f_s)|^2 \cdot |H(f_t)|^2 \cdot |I(f_t)| \cdot |D(f_t)| \quad (5.12)$$

Equation (5.11) is the frequency-dependent response function of the augmented Reichardt correlator model.

Simulating the frequency-dependent response

To simulate the frequency-dependent response we reduce the complexity of the spatio-temporal filter P as follows. In P , three temporal first-order filters are included: H , I , and D . We decrease the number of parameters in the model by replacing the three filters by a single band-pass filter H' .

The filter H' is composed of a low-pass filter $I'(f_t) = H_{LP}(f_t, \tau'_1)$ and a high-pass filter $D'(f_t) = H_{HP}(f_t, \tau'_2)$. In the experiments that follow, we substitute first

$\angle D(f_t) - \angle I(f_t)$ by $\angle D'(f_t) - \angle I'(f_t)$ in equation (5.11) and second $|H(f_t)|^2 \cdot |I(f_t)| \cdot |D(f_t)|$ by $|H'(f_t)|$ in equation (5.12).

Figure 5.3 shows the normalised spatio-temporal frequency response of the augmented Reichardt correlator model so obtained for the following parameter values:

Figure 5.3: The normalised spatio-temporal frequency-dependent response of the augmented Reichardt correlator. Also shown are the spatial and temporal resolution of the correlator, f_s^* and f_t^* , respectively. The correlator's parameter values are summarised in the main text.

$\Delta\varrho_1 = 0.9$ deg, $\Delta\varrho_2 = 4.5$ deg, $\Delta\phi = 1.4$ deg, $\tau'_1 = 0.05$ s, and $\tau'_2 = 0.1$ s. The values were chosen to resemble typical values of flies (see subsection 5.8.1). The response shows a single optimum at $f_s^* = 0.20$ c/deg, $f_t^* = 2.3$ Hz as well as an area of negative response. The values f_s^* and f_t^* are called the tuning frequencies.

The response graph shown in figure 5.3 resembles the frequency-dependent response of biological motion detectors. To illustrate this, we show, in figures 5.4 (a) and (b), the typical spatial and temporal frequency-dependent response of flies. Our model captures two main characteristics of biological motion detectors. First, the graphs in figures 5.3 and 5.4 all show a single optimal value. Second, our model displays negative response for a range of values of spatial resolution. Therefore, we conclude that the simplifications do not reduce the biological validity of our model (cf. Van Santen and Sperling, 1985). In subsection 5.8.1, the frequency-dependent response of the augmented correlator model is compared to the response of biological motion detectors in more detail.

Figure 5.4: The spatial frequency-dependent response of *Drosophila* (a) redrawn from (Buchner, 1976) and the temporal frequency-dependent response of *Calliphora* (b) redrawn from (O’Carroll *et al.*, 1996).

Calculating resolution

To calculate the resolution of our augmented Reichardt correlator model, we proceed as follows. We equate the spatial resolution as the spatial tuning frequencies f_s^* of the frequency-dependent response function (figure 5.3). In analogy, the temporal resolution is equated to f_t^* . Therefore, the resolution can be calculated analytically by finding the zero-crossings of the derivative of equation (5.11). However, due to the complexity of the equation, we calculate the correlator’s resolution numerically. For a given set of parameter values $\Delta\phi$, $\Delta\varrho_{1,2}$, and $\tau'_{1,2}$ we simply calculate the response R_f exhaustively for a range of values of f_s and f_t . From the response values, the values of f_s^* and f_t^* (*i.e.*, the resolution values) are determined.

Setting resolution

In order to optimise the resolution in the experiments that follow, we identify how the different parameters of the augmented Reichardt correlator set the spatial and temporal resolution.

From equations (5.11) and (5.12) it becomes clear that the spatial resolution f_s^* is set by the filter G and the value of $\Delta\phi$. The widths $\Delta\varrho_1$ and $\Delta\varrho_2$ determine the amplitude response of the filter G and, therefore, together with the value of $\Delta\phi$, set the spatial resolution. The temporal resolution f_t^* is determined by the filters H , I , and D , *i.e.*, by H' after applying the substitutions described above. The filter H' is characterised by τ'_1 and τ'_2 which, in turn, set the temporal resolution.

High values of the filter parameters result in low tuning frequencies and *vice*

versa (see equations (5.1) to (5.3)). Therefore, the resolution values are set by varying the parameter values as follows: increasing the filter parameters decreases the resolution, *vice versa*.

Now that we have presented a way to calculate and set our model's resolution, we formulate the model's velocity-dependent response function in subsection 5.2.4.

5.2.4 The velocity-dependent response function

The velocity-dependent response function $R(V_r)$ maps the angular velocity of moving natural images to the output of the augmented Reichardt correlator. The function determines the behaviour of the model of gaze stabilisation as follows. An important factor for the quality of the behaviour is the system gain g (Land, 1992). From figure 5.1 it can be observed that the total gain g of the system equals the product of four constituent gain factors, the correlator gain K_c , the integrator gain K_i , the delay gain K_d , and the variable K :

$$g = K_c \cdot K_i \cdot K_d \cdot K \quad (5.13)$$

In our model, the values of K_i and K_d equal 1. The variable K can be set to a desired value. The velocity-dependent response $R(V_r)$ determines the correlator gain K_c .

$$K_c = R(V_r)/V_r \quad (5.14)$$

To establish the velocity-dependent response function, we employ the frequency-dependent response function (cf. Dror *et al.*, 2001). Any image can be interpreted as a weighted sum of sine-wave gratings of different frequencies that represent the Fourier components of the image (Oppenheim *et al.*, 1997). Strictly speaking, however, calculating the weighted sum of the frequency-dependent response for a range of spatial frequencies does not yield the velocity-dependent response. This effect is due to the non-linear nature of Reichardt correlators. Nevertheless, if we consider the mean velocity-dependent response $\overline{R_v}$, that is averaged over space and time, we find that the response is pseudo-linear, or 'linear in the mean' (Poggio and Reichardt, 1976; Dror *et al.*, 2001).

The mean velocity-dependent response $\overline{R}$ to natural images can therefore be calculated as a sum of individual responses R_f to sine-wave gratings. Equation (5.11) shows that the elements of the sum are weighted by the square of the modulation amplitude ΔI^2 . The value ΔI^2 equals the spectral power $S(f_s)$ of a grating with spatial frequency f_s . The averaged velocity-dependent response to a moving image, therefore, equals a sum of the frequency-dependent responses, weighted by the spatial power spectrum $S(f_s)$ of the image. Consequently, establishing a simple expression for the velocity-dependent response is possible only when $S(f_s)$ is known. From the next paragraph it becomes clear that $S(f_s)$ does not change much with the contents of natural images, which enables us to establish the expression.

Natural images

Statistical analyses of natural images show that the spatial power spectrum $S(f_s)$ of a natural scene is described by the following equation (Ruderman, 1994; Tolhurst and Tadmor, 1997).

$$S(f_s) = S_0 \cdot f_s^{-\alpha} \quad (5.15)$$

where S_0 and α are constants. The weight function $S(f_s)$ varies little with the contents of natural images (Field, 1987; Van Hateren, 1993). For 2-dimensional power spectra, the best-fitting value of α is 2; for 1-dimensional spectra the value is 1 (Van Hateren, 1997; Dror *et al.*, 2001). Since the augmented Reichardt correlator senses motion in the horizontal direction only, we adopt $\alpha = 1$. The value of S_0 is determined by the total energy in the image, which is set by the contrast and the mean illumination of the scene (see subsection 5.5.2 and Van der Schaaf and Van Hateren, 1996).

Since the power spectrum has a similar shape in different natural images, the velocity-response function of Reichardt correlators, varies little with image content as well (Dror *et al.*, 2001). Below we determine the response characteristic of the augmented Reichardt correlator.

The response function

Using the property of pseudo-linearity as explained above, the velocity-dependent response function of our model, $\bar{R}$, is calculated by integrating $R_f(f_s, f_t, \Delta I)$ (equation (5.11)) over spatial frequency. To do this, we apply two substitutions to the frequency-dependent response function. First, we substitute the temporal frequency, according to the definition, by $f_t = V \cdot f_s$. Second, as argued above, the sine wave's squared modulation is substituted by $\Delta I^2 = S(f_s)$, with $\alpha = 1$. The mean velocity-dependent response function $\bar{R}(V)$ of the augmented Reichardt correlator is thus expressed as:

$$\bar{R}(V) = \int_{-\infty}^{\infty} R_f(f_s, v \cdot f_s, \sqrt{S(f_s)}) df_s \quad (5.16)$$

We approximate the response values numerically, *i.e.*, we sum the response R_f of a discrete set of spatial frequencies. The set is bounded such that only response values exceeding 1% of the maximum response are added to the total.

Figure 5.5 shows the velocity-dependent response $\bar{R}(V)$ of the augmented Reichardt correlator for the parameter values summarised in subsection 5.2.3 ($\Delta\theta_1 = 0.9$ deg, $\Delta\theta_2 = 4.5$ deg, $\Delta\phi = 1.4$ deg, $\tau'_1 = 0.05$ s, $\tau'_2 = 0.1$ s, $\alpha = 1$, and $S_0 = 1$). The curve $\bar{R}(V)$ in the figure is bell-shaped when plotted on a logarithmic horizontal scale. The bell shape is also a distinctive feature of the response of biological motion detectors (Kern *et al.*, 2001b). Therefore, the model captures the

Figure 5.5: The velocity-dependent response of the augmented Reichardt correlator to moving natural scenes.

main features of both the frequency-dependent response (subsection 5.2.3) and the velocity-dependent response.

The input velocity for which the correlator is maximally responsive, V^* , is termed the tuning velocity ($V^* = 14.5$ deg/s for the parameter values used here). The motion detector, therefore, responds selectively to input velocities. The tuning velocity determines the motion sensitivity of the detector.

5.2.5 Section summary

Above, we presented our situated model of gaze stabilisation. The model is an elaborated version of a generic model based on a negative feedback control system (Land, 1992). To make the model situated and to represent spatial and temporal resolution in the model, we added a motion detector, a low-pass filter, and a source of noise (figure 5.1). The additions allow us to study the trade-off between spatial and temporal resolution by optimising visuo-motor behaviour, which was not possible with the original model.

The most important component of the model is the augmented Reichardt correlator that was mathematically described and characterised by its frequency-dependent and velocity-dependent response functions. We showed that the augmented correlator captures the most important characteristics of biological motion detectors. Therefore, we presuppose that our model is a plausible representation of biological gaze-stabilisation behaviour. A full evaluation of the model follows in section 5.8.

In the following sections, it is investigated how the three selected determinants

— sensitivity to motion, light, and contrast — set the optimal resolution of the model, *i.e.*, the resolution that produces the optimal gaze-stabilisation behaviour (see section 5.1). We start, in the following section, by studying how the resolution sets the motion sensitivity.

5.3 How visual resolution sets motion sensitivity

In this section, we determine how visual resolution sets the motion sensitivity of our model. Establishing this relation is the first step in understanding whether the determinant motion sensitivity causes a trade-off between spatial and temporal resolution (section 5.1.1).

The relation between resolution and motion sensitivity is generally assumed to be expressed as (cf. O'Carroll *et al.*, 1996):

$$V^* = a \cdot \frac{f_t^*}{f_s^*} \quad (5.17)$$

with V^* representing the tuning velocity (motion sensitivity), f_s^* and f_t^* as spatial and temporal resolution, respectively, and $a = 1$ (*i.e.*, the definition of temporal frequency; see subsection 5.2.3). We argue that in general $a > 1$, *i.e.*, the motion sensitivity is larger than the fraction of temporal resolution over spatial resolution.

In subsection 5.3.1 we show one specific example in which the value of a exceeds unity. Thereafter, in subsections 5.3.2 to 5.3.4 we discuss the underlying causes and the range of values that the variable a can take. Finally, in subsection 5.3.5 the section is summarised.

5.3.1 The relation in the model

To determine the relation, *i.e.*, the value of a in our model, we use equations (5.11) and (5.16) to calculate the variables V^* , f_s^* , and f_t^* . The three variables are calculated for a biologically plausible range of values of the model parameters. The parameter $\Delta\varrho_1$ is assigned a geometric sequence of values between 0.09 and 9. To τ_1 we assign a sequence between 0.005 and 0.25. Moreover, $\Delta\varrho_2 = 5 \cdot \Delta\varrho_1$; $\Delta\phi = 1.6 \cdot \Delta\varrho_1$; and $\tau_2 = 2 \cdot \tau_1$. The plausibility of the values is demonstrated in section 5.8. Figure 5.6 shows the obtained values of the tuning velocity (V^*) as a function of the spatial and temporal tuning frequency (f_s^* and f_t^* , respectively).

To show how the resolution sets the motion sensitivity, we fit the obtained values to equation (5.17) as follows. First, the data points are interpolated to obtain N iso-metric lines that correspond to N values of V^* ; second, for every line, we calculate the mean fraction f_t^*/f_s^* ; and third, we fit the values of V^* to the calculated fractions through least-squares analysis. For the particular values of the variables summarised above and $N = 10$, we obtain $a = 1.25$. The mean relative error between the approximated values of V^* and the measured data shown in figure 5.6 equals 5%.

Figure 5.6: Velocity tuning of the augmented correlator as a function of frequency tuning on logarithmic scales.

Clearly the value of a obtained here is larger than the value generally used ($a = 1$). Below we study why a does not equal unity in our model. For this we investigate two important features of natural stimuli and of biological motion detectors that are captured in our model. First, we investigate the statistics of natural images (subsection 5.3.2). Second, we study the negative parts in the frequency-dependent response (subsection 5.3.3). The results show that the two features modify the tuning frequencies of the model and of biological motion detectors. In subsection 5.3.4 we show to what extent the tuning frequencies are modified, *i.e.*, the range of possible values that the variable a can take.

5.3.2 The role of spatial natural image statistics

Statistically, the low spatial frequencies in natural images have a higher power than high frequencies (equation (5.15)). We investigate whether this distribution results in a value $a > 1$.

To determine the role of spatial natural image statistics in setting the value of a , all factors in equation (5.11) that depend on temporal resolution, as well as the modulation ΔI are normalised to one yielding $R_f = |G(f_s)|^2 \cdot \sin(2 \cdot \pi \cdot f_s \Delta \phi)$. Figure 5.7 shows the equation with the same parameter values as used in subsection 5.2.3. The lower curve in the figure shows the spatial frequency-dependent response $R_f(f_s)$ to sine-wave grating. As in subsection 5.2.3, the response is maximal at the tuning frequency $f_s^* = 0.20$ deg/s. To determine the tuning frequency f' for

Figure 5.7: Response as a function of spatial frequency.

natural images we calculate the frequency response weighted by the spatial power spectrum: R_t/f_s (upper curve). Doing so we find that the role of spatial natural image statistics is to decrease the tuning frequency to $f' = 0.17\text{deg/s}$. By definition, we expect that the tuning velocity for moving natural images is expressed as $V^* = f_t^*/f' = f_s^*/f' \cdot f_t^*/f_s^*$. When substituting $f_s^*/f' = 0.20/0.17 = 1.18$ we arrive at equation (5.17) with $a = 1.18$. Therefore, in this example the spatial image statistics set the constant a to a value larger than unity. Due to the $1/f_s$ shape of the power spectrum, the difference between f_s^* and f' becomes larger when the width of the curves increases. Therefore, we expect that the value of a generally depends on the width of the curves in figure 5.7 (*i.e.*, the width of the spatial band-pass filter G).

We showed an example in which a is greater than 1. However, when using the same pass-band widths as before, the value of a found here does not correspond to the value $a = 1.25$ determined in subsection 5.3.1. Therefore, before studying the role of band-pass width in subsection 5.3.4, we investigate the role of negative response, another factor that plays a role in the relation between the resolution and motion sensitivity.

5.3.3 The role of negative response

The second explanation of the value of a found in subsection 5.3.1 concerns the role of the negative part of the frequency-dependent response. The role of negative response is explained using a contour plot of the typical frequency-dependent response of the augmented Reichardt correlator shown in figure 5.8. In the figure, the maximal response (+) is located at the tuning frequencies f_s^* and f_t^* . The minimal (negative)

Figure 5.8: Contour plot of the typical frequency-dependent response of the augmented Reichardt correlator model. Shown are the maximal (+) and minimal (-) response, as well as the contours corresponding to 50% of the minimal/maximal response.

response is indicated by a minus sign (-). Around the extremes, the contours of 50% response values are drawn as ellipses. The diagonal dashed lines represent the equation $f_t = V \cdot f_s$ for different background velocities V . Both axes have a logarithmic scale, which explains why the dashed lines run parallel to each other.

To understand how the area of negative response influences the tuning velocity V^* , we compare the responses to three velocity values. For this we first consider the line λ_0 that passes the point (f_s^*, f_t^*) and represents a velocity $V_{\lambda_0} = f_t^*/f_s^*$. We also consider two other lines, λ_1 and λ_2 , that are translated versions of λ_0 and represent velocities $V_{\lambda_1} < V_{\lambda_0} < V_{\lambda_2}$. In figure 5.9, we show the cross sections $S(V, f_s)$ of the frequency-dependent response along the three lines λ_0 , λ_1 , and λ_2 . Following equation (5.16) we calculate the velocity-dependent response by integrating the cross-sectioned frequency-dependent response $S(V, f_s)$. More specific, to determine the isolated role of the area of negative response we integrate the *unweighted* response: $R(V_\lambda) = \int S(V_\lambda, f_s) df_s$. Comparing the calculated values yields $R(V_{\lambda_2}) > R(V_{\lambda_0}) > R(V_{\lambda_1})$. Since the response is maximal at V^* we deduce from the comparison that $V^* > V_{\lambda_0}$. Above, we defined V_{λ_0} as the fraction of the temporal over spatial tuning frequencies. Therefore, the role of the area of negative response in the relation between resolution and motion sensitivity is that the motion sensitivity V^* exceeds the fraction of the resolution values f_t^*/f_s^* , i.e., the variable a in equation (5.17) is greater than unity.

The response $R(V_{\lambda_2})$ exceeds $R(V_{\lambda_0})$ since the line λ_2 does not cross the 50%

Figure 5.9: Cross section of the frequency-dependent response along the lines λ_0 , λ_1 , and λ_2 that are drawn in figure 5.8.

area of negative response. Therefore, we expect that the exact position of the line λ' with $V_{\lambda'} = V^*$, depends on the widths of the ellipses in figure 5.8 which are set by the pass-band widths of the spatial and temporal band-pass filters. In the next subsection, we measure the influence of changing the spatial and temporal pass-band widths on the value of a . Doing so, we verify the explanations for the value of a given in subsections 5.3.2 and 5.3.3.

5.3.4 The role of pass-band width

The findings from the two previous subsections bear consequences for the role of the filter pass-band width in determining the value of a . In this subsection we verify the consequences experimentally.

According to subsection 5.3.2 the spatial tuning frequency of the frequency-dependent response, weighted by the spatial power spectrum, is lower than the tuning frequency of the non-weighted response. Widening the pass band of the spatial pre-filter makes the system more responsive to both high and low frequencies. As demonstrated by the spatial power spectrum of natural images (equation (5.15)), the low frequencies have more power than the high frequencies. Therefore, the spatial tuning frequency is decreased. As a result, the value of a is expected to increase when widening the spatial pass band. Similarly, according to subsection 5.3.3, widening the temporal pass band is also expected to increase the value of a , due to the influence of the area of negative response in the frequency-dependent response.

Varying pass-band width

To simulate different pass-band widths we vary the ratios of the spatial and temporal filter constants over the following ranges: $\frac{\Delta\theta_2}{\Delta\theta_1} \in [5, 40]$, homogeneously spaced, and $\frac{\tau_2}{\tau_1} \in [2, 5]$, also homogeneously spaced. We keep the fraction $\frac{\Delta\theta_2}{\Delta\theta_1}$ constant at a value 1.6 (see subsection 5.2.3). The values are chosen such as to demonstrate the effects of varying the pass-band width best. When varying the spatial pass-band width, we keep the temporal pass-band width at a constant base value, *vice versa*. The base values are the same as used in subsection 5.3.1: $\frac{\Delta\theta_2}{\Delta\theta_1} = 5$ and $\frac{\tau_2}{\tau_1} = 2$. For every ratio used, we calculate the value of a and the width of the frequency-dependent response surface. The value of a is calculated by fitting equation (5.17) as was explained in subsection 5.3.1. The spatial width of the response surface is measured as follows. For a constant temporal frequency $f_t = f_t^*$ we find the spatial cut-off frequencies f_1 and f_2 where the response equals 50% of the maximal response ($f_2 > f_1$). The spatial width Δf_s^* is expressed in octaves: $\Delta f_s^* = \log(f_2/f_1)/\log(2)$. The temporal width Δf_t^* is calculated using the same procedure, with $f_s = f_s^*$ fixed and f_1 and f_2 representing the temporal cut-off frequencies.

Results

In figure 5.10 we draw the calculated values of a as a function of spatial and temporal pass-band width. In the figure, it is shown that the value of a grows when either

Figure 5.10: The value of a in equation (5.17) as a function of spatial and temporal pass-band width.

the spatial or the temporal pass band is widened. The relative error in fitting equation (5.17), averaged over all spatial pass-band widths, equals 5%; the mean error over all temporal pass-band widths equals 4%. The results show that the widths of the pass bands of the spatial and temporal filters affect the value of the constant a in equation (5.17).

In conclusion, figure 5.10 shows that the constant a can have a range of values. The value $a = 1.25$ found in subsection 5.3.1 is determined by the ratios $\frac{\Delta\varphi_2}{\Delta\varphi_1}$ and $\frac{\tau_2}{\tau_1}$ employed. Hence, the role of the spatial natural image statistics was verified and the role of the area of negative response in the relation between resolution and motion sensitivity was confirmed. Owing to both roles the value of a in equation (5.17) is set to a value greater than unity.

5.3.5 Section conclusion

In conclusion, we state that the tuning velocity V^* of the augmented Reichardt correlator is proportional to the fraction of temporal over spatial tuning frequency f_t^*/f_s^* (equation (5.17)). In other words, the visual resolution sets the motion sensitivity to a value proportional to the fraction of the resolution values.

Moreover, it was determined that the constant of proportionality a exceeds 1 and depends on the pre-filters' pass-band widths: widening the pass bands leads to larger values of a . We found that the two factors that play a role in this dependency are: the spatial natural image statistics (*i.e.*, the power density function) and the presence of a negative region in the frequency-dependent response. Therefore, our results lead us to a second conclusion which states that the motion sensitivity of our model is set by the visual resolution as well as by the width of the frequency-dependent response of the augmented Reichardt correlator.

5.4 The optimal motion sensitivity

We optimise the model's motion sensitivity to a value that optimises the behaviour of the model. The optimal value is subsequently used to determine whether the determinant motion sensitivity causes a trade-off between the spatial and temporal resolution of our model.

By definition, the velocity of moving images sets the optimal motion sensitivity. Therefore, we optimise the motion sensitivity as a function of the velocity of moving images. The motion sensitivity of the model is represented by the tuning velocity V^* of the augmented Reichardt correlator (subsection 5.2.4). Below we show that an optimal tuning velocity optimises the model behaviour in three ways: first, by optimising the signal-to-noise ratio (SNR) of the correlator's output signal; second, by optimising the gain control of the gaze-stabilisation model; third by optimising the SNR and gain control jointly. Each way is studied in subsections 5.4.1, 5.4.2, and 5.4.3, respectively. Thereafter, we provide a summary in subsection 5.4.4.

5.4.1 Signal-to-noise ratio optimisation

We aim to find the tuning velocity that optimises the signal-to-noise ratio and thus maximises the success of gaze-stabilisation behaviour. As was shown in subsection 5.2.4, the augmented Reichardt correlator responds to a limited range of input velocities. For input values outside this range, the response is so small that it is masked by noise which is omnipresent in neural systems (Pelli, 1990). From the control diagram of gaze-stabilisation behaviour (figure 5.1), we find that a noisy correlator response leads to a large variance of the output signal V_s , *i.e.*, low-quality behaviour. Van Hateren (1993) assumed that the SNR in the early stages of visual processing is optimised with respect to the typical ego-motion of an animal. Following this assumption, we establish the relation between tuning velocity and ego-motion that optimises the SNR in the output signal of the correlator and thus optimises the gaze-stabilisation behaviour.

To find the relation between tuning velocity and ego-motion that leads to an optimal SNR, we will calculate the open-loop response of the model of gaze-stabilisation. Figure 5.11 is based on the control diagram of the model introduced in section 5.2. We added a switch to change between open-loop (open switch) and closed-loop

Figure 5.11: The control diagram for gaze stabilisation behaviour augmented by a switch for open-loop and closed-loop control.

(closed switch) behaviour. In open-loop circumstances, the system receives no feedback from its actions, *i.e.*, $V_r = V_b$. The new setup allows us to find the motion sensitivity that optimises the behaviour through SNR optimisation only.

In our experiments, we set the background velocity V_b to a constant value to simulate ego-motion, *i.e.*, the visual background rotates with constant velocity. To find the optimal tuning velocity, V_{opt}^* , as a function of background velocity, V_b , we measured the behaviour for a range of values of V^* and V_b . Then, the quality of the behaviour was estimated by measuring the time-average and variance of the steady-state response V_s . For every value of V_b , we determined the value of V^* that produces the best open-loop behaviour, *i.e.*, that yields the lowest response variance.

In the next paragraphs we discuss two experimental issues: the settings of model parameters and the type of noise used. Thereafter we present the results of the open-loop experiments.

Model parameters

In the experiments we kept some of the model parameters fixed: the time constant τ of the low-pass filter LP was set to 750 ms; the system latency equalled 75 ms. Both values are biologically plausible (Warzecha and Egelhaaf, 1996). Moreover, the constant S_0 in the spatial power spectrum of natural images (equation (5.15)) was set to 1, *i.e.*, we applied a fixed illumination and contrast. The value of the internal gain constant K was adjusted to the tuning velocity V^* and the background velocity V_b to yield a desired open-loop system gain g (see equation (5.13)). In our model, the steady-state gain K_i of the integrator LP is constant ($K_i = 1$). In contrast, the Reichardt correlator gain K_c is variable and depends on the input velocity V_r and the velocity-dependent response function R (equation (5.14)) which is characterised by V^* . By using constant values of V_b , K_c was fixed for each combination of V_b and V^* (in open-loop systems $V_r = V_b$). The total gain of the system was set to $g = 0.9$ which is in the range of gain values observed in biological systems (cf. Collett, 1980). The value indicates that the system compensated for 90% of the background motion. To determine the internal gain constant K , we measured the correlator gain K_c for the employed values of V_b and V^* , and set $K = g/K_c$.

Noise

In biological gaze-stabilisation systems, there exist various sources of noise. At low light levels, photon shot noise (Snyder, 1979), spontaneous thermal activation (dark light; Barlow, 1956), and photon transduction noise are the most prominent sources in the receptor system (Van Hateren, 1993; Warrant, 1999). Further up in the neural system, synapses and the limited dynamic range of neural channels add more noise (Van Hateren, 1993). Finally, the clock-firing of motor neurons causes a tremor in the muscles adding more noise to the system (Eizenman, Hallett, and Frecker, 1985).

In our experiments, we restricted ourselves to modelling the most limiting source of noise, *i.e.*, the type of noise that is most disturbing to the behaviour. Lewen *et al.* (2001) measured the neural signals of flies using naturalistic motion stimuli. They found that photon shot noise is limiting the accuracy of the neural response, rather than neural noise. The animals investigated were maintained in open-loop circumstances, as in our experimental setup. Therefore, in our experiments we applied photon shot noise that is best modelled by drawing a variable from a Poisson distribution (Snyder, 1979). As in subsection 4.2.2, we approximated the Poisson distribution as a Gaussian with a standard deviation equal to the square root of the mean.

Results

Using the experimental setup described above, we determined the optimal tuning velocity V_{opt} for a range of values of V_b . In figures 5.12 (a) to (d) we show the mean and standard deviation of the response for four values of V_b and 20 values of V^* per V_b value. In the figures, the upper curves depict the sum of mean and standard

Figure 5.12: Open-loop gaze-stabilisation behaviour. In every graph, the mean response V_s (lower curves) and the sum of mean and standard deviation of the response (upper curves) are shown as a function of tuning velocity V^* . The scales of the four graphs differ to show the qualitative differences between the curves in different graphs.

deviation of the response, while the lower curves represent the mean only. The upper curves all have a minimal value, *i.e.*, the response shows the least variance, at $V^* = V_{opt}^*$. Based on the figures, we expect that the optimal tuning velocity V_{opt}^* increases when the background velocity V_b increases. In figure 5.13, we plot V_{opt}^* as a function of V_b . We fit the V_{opt}^* values to the identity function (dashed line). The fit has a mean relative error of 11%, which is rather large. However, the fitted curve lies within the error bounds of most data points, which makes the fit acceptable.

Figure 5.13: The optimal tuning velocity V_{opt}^* drawn as a function of the background velocity V_b in open-loop control.

Summary

In summary, we aimed to find the tuning velocity that optimises the SNR and thus maximises the success of gaze-stabilisation behaviour. The optimal tuning velocity, V_{opt}^* , is set by the background velocity V_b according to equation (5.18)

$$V_{opt}^* = c \cdot V_r = c \cdot (1 - G) \cdot V_b \quad (5.18)$$

with the closed-loop gain G equal to 0 (open-loop control) and c equal to 1, *i.e.*, the identity function.

The tuning velocity V^* is proportional to the fraction of temporal resolution over spatial resolution f_t^*/f_s^* (equation (5.17)). Therefore, equation (5.18) also shows that the optimal fraction of resolution values is determined by V_b to optimise the SNR in the system.

5.4.2 Gain-control optimisation

As outlined in the beginning of the section, the second way in which the tuning velocity may maximise the success of gaze-stabilisation behaviour is by optimising the gain control. Closed-loop feedback systems such as the gaze-stabilisation system are prone to instability (Oppenheim *et al.*, 1997). Variations in the system's response are caused by different factors. High values of the gain, the noise level (cf. subsection 5.4.1), and the latency of the system all lead to instability (Land,

1992; Feng and Plamondon, 2001). The instability is expressed as high-amplitude jitter: large oscillations in the system's response velocity that reduce the quality of visual perception. Warzecha and Egelhaaf (1996) showed that Reichardt correlators have a built-in gain-control mechanism that reduces the jitter. The level of gain control is set by the value of the tuning velocity V^* , i.e., by the motion sensitivity. The aim of this section is to understand how the optimal tuning velocity V_{opt}^* is related to the velocity V_b of visual motion to optimise the gain control of our model of gaze stabilisation. In the following, we continue by explaining how gain control is accomplished in the augmented Reichardt correlator. Then we provide experimental details to determine the relation between V_{opt}^* and V_b , after which the results are presented and discussed.

Gain control in the augmented Reichardt correlator

Warzecha and Egelhaaf (1996) showed that Reichardt correlators have intrinsic gain-control properties. In this paragraph we determine the gain control that is displayed by our augmented Reichardt correlator model. Figure 5.14 shows the gain K_c (equation (5.14)) of four correlators with increasing tuning velocities V_1^* (top curve) to V_4^* (bottom curve). For retinal velocities V_r above a threshold value $\theta(V^*)$ the gain is attenuated. As a result, large input velocities, caused by heavy jitter, are transmitted at low gain, which reduces the response variance. Therefore, we find that the

Figure 5.14: The gain of four augmented Reichardt correlators with different tuning velocities as a function of retinal velocity V_r . The variable $\theta(V^*)$ denotes the threshold velocity above which the gain is controlled by attenuation.

augmented Reichardt correlator has intrinsic gain-control properties. As indicated in the figure, the threshold $\theta(V^*)$ of each correlator is set by the tuning velocity V^* : the larger V^* , the larger $\theta(V^*)$. Hence, the tuning velocity sets the level of gain control and, therefore, controls the stability of the model of gaze stabilisation. In general, a model that uses the lowest tuning velocity employs maximal gain control and thus produces the stablest behaviour.

Below, we first provide the details of our experiments: noise, correlator response, and gain tuning. Thereafter, we present and discuss the results.

Noise

To find the values of V_{opt}^* that exclusively optimise the gain control of the augmented Reichardt correlator we used Gaussian noise. The standard deviation σ of the noise equalled a fixed fraction f of the output signal R_i of the integrator LP : $\sigma = f \cdot R_i$. In this way the SNR of the noisy integrator output constant did not depend on the input velocity, *i.e.*, we optimised gain control, not the SNR (cf. subsection 5.4.1).

Natural visuo-motor systems display significant variance in their response over time (Eizenman *et al.*, 1985). The noise level in our model was set such as to resemble natural gaze-stabilisation behaviour best. Warzecha and Egelhaaf (1996) found that in flies the standard deviation of retinal velocity V_r equals as much as eight times the mean. In our experiments we set $f = 3$, which resulted in a similar response variance. Due to the large relative standard deviation, we measured the retinal slip velocity of each controller during $1 \cdot 10^6$ simulation steps. Pilot studies showed that this number guarantees a reliable estimate of the system's behaviour.

Now that we have introduced the type and amount of noise we proceed by showing how the gain parameter K was adjusted to the two variables in the experiment, namely to the background velocity V_b and the tuning velocity V^* .

Gain tuning

To find the motion sensitivity V_{opt}^* that optimises gain control, we tuned the gain factor K in the model of gaze stabilisation (see figure 5.11) to a target value K^* . The procedure is as follows. As was shown above, gain control depends on the values of V_r and V^* (cf. figure 5.14). Therefore, to determine V_{opt}^* we kept V_r constant at a target value V_r^* . In closed-loop control, the mean value of V_r depends on V_b and the system gain g (subsection 5.2.4), *i.e.*, on the value of K (equation (5.13)). Therefore, in order to find V_{opt}^* , the gain factor K was tuned to a value K^* to yield constant V_r^* values.

We tuned the gain as follows. We first set K to an initial value K_{init} and subsequently automatically updated the value to find K^* that yielded the desired retinal velocity V_r^* . Below we discuss the initialisation and update rule, including a test for convergence.

Gain initialisation To calculate the initial value K_{init} three equations were combined into one. First, we used equation (5.13) which expresses the open-loop gain g as the product of K and other gain factors. Second, the definition $g = V_s/V_r$ was adopted. Third, we employed the steady-state equation of the system $V_s = V_b - V_r$. With $K_i = K_d = 1$ (subsection 5.2.4) the equation that yields K_{init} as a function of V_b and V_r^* reads:

$$K_{\text{init}}(V_b, V_r^*) = \frac{g}{K_c} = \frac{V_s}{V_r^*} \cdot \frac{1}{K_c} = \left(\frac{V_b}{V_r^*} - 1 \right) \cdot \frac{1}{K_c} \quad (5.19)$$

Although equation (5.19) provided good initialisations of K , we found that K_{init} was unequal to the value of the target value K^* . To illustrate this fact, we measured the mean steady-state retinal slip velocity $\bar{V}_r$ for a range of values of K . Using the model parameter values (summarised in subsection 5.4.1), Gaussian noise, and fixed values of V_b and V^* , we obtained results that are shown in figure 5.15. Subsequently, after

Figure 5.15: The mean retinal slip velocity $\bar{V}_r$ of the model of gaze stabilisation measured for a range of values of the gain factor K .

setting the target gain and velocity (K^* and V_r^* , respectively) to arbitrary values, we calculated K_{init} using equation (5.19). In the figure, we also indicate K^* and V_r^* , as well as the calculated value K_{init} and the corresponding velocity $V_{r,\text{init}}$. It is shown that $V_{r,\text{init}}$ is larger than the target value V_r^* . Moreover, K_{init} is smaller than K^* . Therefore, we adapted the initial value K_{init} somewhat to obtain the target value K^* .

To investigate why equation (5.19) does not produce the target gain value K^* , we measured the model behaviour using three other arbitrary correlators. With $K = K_{\text{init}}$, the actual retinal slip velocity $\bar{V}_r$ is drawn as a function of the target

Figure 5.16: The mean retinal slip velocity $\overline{V}_r$ of three versions of the model of gaze stabilisation. The value of $\overline{V}_r$ was measured for a range of values of the target retinal slip velocity V_r^* .

velocity V_r^* in figure 5.16. It is shown that the curves representing high values of V^* approached the identity function (dashed line) better than the curves associated with low values of V^* . Correspondingly, we found that the difference between K_{init} and K^* decreased when V^* increases. This fact is explained by the gain control properties of the correlators: using low tuning velocity generally results in much gain control and a low closed-loop gain. As a result, the actual retinal slip velocity appeared to be lower than expected from equation (5.19). In the following section we introduce an iterative update rule to tune the gain factor K .

Iterative update rule To adapt the K -value automatically, we employed an iterative update rule inspired by artificial neural network training (Hertz *et al.*, 1991; Mitchell, 1997). The update rule is described as follows. For a fixed value of V_b , we set a target retinal slip velocity V_r^* and initialise K according to equation (5.19). We then calculate the target open-loop gain $g^* = V_s/V_r^* = V_b/V_r^* - 1$. Subsequently, the behaviour of the model is measured and the actual value of $g = \overline{V}_s/\overline{V}_r$ is calculated for the current value of K . Finally, we determine the error e between the target gain and the actual gain value:

$$e = g^* - g \quad (5.20)$$

To calculate the part of the error, e_K , that is due to the current K value, we divide e by the correlator gain K_c and rewrite the expression using equation (5.13):

$$e_K = \frac{e}{K_c} = \frac{e}{g} \cdot K \quad (5.21)$$

with $K_i = K_d = 1$ as shown above.

After calculating e_K , we update K by adding a value ΔK in order to minimise the difference between the actual gain and target gain. The value ΔK is calculated as:

$$\Delta K = \eta \cdot e_K = \eta \cdot \frac{e}{g} \cdot K \quad (5.22)$$

In the equation, η represents the learning parameter.

Stability of the update rule The update rule is stable, *i.e.*, the error converges to a fixed value $e = 0$, only when the parameter η is kept in bounds. In appendix B we calculate the bounding conditions. Based on the calculations we adopted $\eta = 1$ in our experiments. However, in the appendix we also show that for some values of V^* and V_r^* , the error cannot converge. To overcome this problem, the minimal error value is stored during updating and the update process is terminated when K exceeds a threshold value θ_K . After termination, the K -value associated to the minimal error is returned.

Now that we have described the iterative gain update rule, we proceed by explaining how the correlator response is calculated in the experiments.

Correlator response

We calculated the correlator response to incoming velocities by consulting a lookup table¹. The lookup table was constructed in advance and consisted of the correlator gain values K_c (equation (5.14)) for 8000 values of input velocity V_r (both positive and negative). The set of V_r values was exponentially spaced to obtain a homogeneous spread of response values (cf. figure 5.14). Moreover, the minimal and maximal V_r values were chosen such that the minimal gain value equalled less than 1 % of the maximal gain. For input values outside the V_r range we assigned a gain $K_c = 0$.

Results

In this paragraph we present our results. To determine the optimal tuning velocity as a function of background velocity, we measured the closed-loop response of

¹The lookup table was constructed for reasons of time efficiency. A quick-search algorithm was used to consult the lookup table using. We estimated that the use of lookup tables reduced the calculation time ten fold as compared to calculating and averaging the response of a large number of Reichardt correlators. Using Pentium III, 1GHz computers, the time to execute a statistically significant number of experiments was thus reduced to the order of days instead of weeks.

30 versions of the model (each version was assigned a different tuning velocity V^*) to 10 different values of the background velocity V_b . For every combination of V^* and V_b , we tuned the value of K , as described above, to obtain a target closed-loop gain $G^* = 0.9$ (cf. Collett, 1980). The time constant τ of the integrating filter LP was set to 750 ms, the latency to 75 ms, and S_0 to 1 (see subsection 5.4.1). Figures 5.17 (a) to (d) show the results. The graphs show the time-averaged steady-state retinal

(a) $V_b = 30.0$ deg/s (b) $V_b = 45.6$ deg/s (c) $V_b = 69.4$ deg/s (d) $V_b = 105.5$ deg/s

Figure 5.17: The mean retinal velocity (lower curves) and the sum of the mean and standard deviation of the retinal velocity as a function of the tuning velocity for four different background velocities.

velocities $\bar{V}_r$ (lower curves) and the sum of the mean and standard deviation of the retinal velocity (upper curves) measured for a range of 30 tuning velocities V^* and four background velocities V_b . Below we show that the results indicate that the optimal tuning velocity has a lower and an upper bound. Moreover, it is explained how the bounds depend on the background velocity V_b , and the target gain G^* .

Thereafter, we determine the tuning velocity that optimises the gaze-stabilisation behaviour by optimising gain control.

Lower bound It is shown in the figures that the mean retinal slip velocity of correlators that have low tuning velocities is higher than of correlators that have high tuning velocities. This observation is explained below by referring to the gain-control properties of the correlator. In figure 5.18 we draw the closed-loop gain, calculated from the data shown in figure 5.17 (a) to (d), as a function of V^* . We found that, although we tuned the gain to a target value, the theoretically maximal

Figure 5.18: Maximal closed-loop gain of 30 correlators for four different background velocities.

gain is lower than the target gain when using low tuning velocities.

The results indicate that the optimal tuning velocity V_{opt}^* has a lower bound. We define the lower bound as the smallest value of V^* where the closed-loop gain equals the target gain G^* . In figure 5.18, it is shown that the curves describing the closed-loop gain have an increasing part and a constant part. The lower bound is indicated to lie at the point of intersection of the increasing and constant parts of the curves. From figure 5.18 it is derived that the point of intersection, *i.e.*, the lower bound of V_{opt}^* depends on two variables, *viz.* the target gain G^* and the background velocity V_b (it is shown in the figure that the curves are shifted rightward when V_b increases). Altogether, we find that the tuning velocity has a lower bound that depends on the target gain G^* and the background velocity V_b .

Upper bound In the figures 5.17 (a) to (d), it is also shown that the standard deviation of the retinal slip velocity, relative to the mean, is small for correlators with

low tuning velocities. As determined above, the low standard deviation is due to high gain control. In contrast, we found that correlators with high tuning velocities generally display a much larger standard deviation because the latter correlators employ little gain control (cf. figure 5.14). A large standard deviation indicates much jitter, *i.e.*, poor gaze-stabilisation behaviour. Therefore, the results show that high tuning velocities optimise the behaviour much less than low tuning velocities. Consequently, setting a maximal level of jitter yields an upper bound to the optimal tuning velocity V_{opt}^* .

The curve that represents the sum of the mean and standard deviation of the retinal slip velocity is shifted rightward when the background velocity is increased (compare figures 5.17 (a) to (d)). Therefore, it is concluded that the tuning velocity V_{opt}^* has an upper bound that increases with increasing background velocity V_b .

The tuning velocity that optimises gain control We define the optimal tuning velocity as the tuning velocity V^* of the augmented Reichardt correlator that (1) yields the desired closed-loop gain G^* and (2) produces the smallest standard deviation of the retinal velocity, *i.e.*, the least jitter. We determined the optimal tuning velocity V_{opt}^* for the range of values of the background velocity used. Based on the properties of the lower and upper bound outlined above, we expected that the optimal tuning velocity V_{opt}^* depends on both the target gain value G^* and the background velocity V_b , *i.e.*, on the retinal slip velocity ($V_r = (1 - G) \cdot V_b$). Therefore, the optimal tuning velocity is drawn in figure 5.19 as a function of retinal slip velocity. In the figure, we find that the optimal tuning velocity grows as a function

Figure 5.19: Optimal tuning velocity as a function of the background velocity in closed-loop control.

of the retinal slip velocity. The dotted line represents equation (5.18) that is fitted to the data yielding a coefficient $c = 23.1$. The mean relative error between the fitted curve and the data is 7%, which is comparable to the error in fitting the results from the open-loop experiments (see figure 5.13). Altogether, we find that the optimal tuning velocity depends on the retinal slip velocity according to equation (5.18).

Summary

In summary, we determined the optimal tuning velocity that maximises the success of gaze-stabilisation behaviour by optimising the gain control. The optimal tuning velocity was determined as a function of the retinal slip velocity according to equation (5.18). When comparing the results of this subsection to those of subsection 5.4.1 we find that the tuning velocities that yield optimal gain control are much higher ($c = 23.1$) than the tuning velocities yielding optimal SNR ($c = 1$).

So far, we have studied two ways in which the tuning velocity optimises the gaze-stabilisation behaviour. First by optimising the SNR (subsection 5.4.1) and second by optimising the gain control (subsection 5.4.2). In the following subsection, we study the third way: by combining the optimisation of SNR and gain control.

5.4.3 Combined SNR and gain-control optimisation

Biological systems maximise the success of their gaze-stabilisation behaviour by optimising both the SNR and gain control (Warrant, 1999; Warzecha and Egelhaaf, 1996). In this subsection we emulate the biological situation in our model, *i.e.*, we study how the tuning velocity optimises the gaze-stabilisation behaviour by optimising the SNR and gain control jointly.

In the previous two sections, we found different values for the coefficient c when optimising either the factor SNR or the factor gain control. To study the joint optimisation of the two factors, we investigate which factor is limiting and what value of c is optimal. As was mentioned before, the noise level determines the stability of the system and, therefore, the amount of gain control necessary. Consequently, we expect that the noise level determines which optimisation factor is limiting. To verify our expectations, we performed a series of experiments, in each of which we applied a different noise level.

The experiments were performed in the same way as described in subsection 5.4.2. However, to emulate the biological situation, we merged the experimental set-ups from the two previous subsections. We applied input-dependent, Poisson-distributed noise as in subsection 5.4.1. Moreover, we employed closed-loop control (cf. figure 5.11, switch closed) as was done in subsection 5.4.2. Below, we show how the noise level was varied in the system. Thereafter we present the results of the experiment followed by an analysis of the limiting optimising factors.

Noise level

As before, we simulated Poisson-distributed noise by drawing the noisy signal R' from a Gaussian distribution $GD(\mu, \sigma)$ with mean $\mu = R$ and standard deviation $\sigma = \sqrt{R}$ (R represents the original signal). Setting the noise level is described as follows. Before drawing a number from GD , the signal R is divided by a factor M (i.e., $\mu = R/M$, $\sigma = \sqrt{R/M}$). The number taken from the distribution is subsequently multiplied by the same factor M , to yield the value R' :

$$R' = M \cdot GD\left(\frac{R}{M}, \sqrt{\frac{R}{M}}\right) \cong M \cdot \left(\frac{R}{M} \pm \sqrt{\frac{R}{M}}\right) = R \pm \sqrt{M \cdot R} \quad (5.23)$$

From equation (5.23), it follows that the variance of the noisy signal R' is determined by the factor $\sqrt{M \cdot R}$, i.e., by the noise level M . The mean of the noisy signal $\bar{R}' = R$ is not affected by the value of M .

Results

Figure 5.20 shows the results of the experiments. In the figure, three curves indicate the optimal tuning velocity V_{opt}^* as a function of the retinal slip velocity V_r that optimises both the SNR and the gain control of the system². Every curve corresponds to a different noise level $M \in \{1.3, 2.5, 5\}$. To find the relation between the optimal

Figure 5.20: Optimal tuning velocity as a function of the input velocity V_r to optimise both the signal-to-noise ratio and system gain control.

²We removed the error bars in order to increase readability.

tuning velocity and the retinal slip velocity we fitted equation (5.18) to the measured values. The fitted curves are drawn in the figure as dashed lines. The coefficients have values $c = 9.5$, $c = 4.6$, and $c = 2.1$, in top-to-bottom order.

Limiting optimising factors

The bottom-most curve (low noise level) approaches the curve shown in figure 5.13 in which only the SNR was optimised ($c = 1$). In contrast, the top-most curve (high noise level) is closer to the curve shown in figure 5.19 in which only the gain control of the system was optimised ($c = 23.1$). Therefore, the results suggest that the noise level M determines the optimal value of c . Moreover, they indicate to what extent each optimisation factor — SNR optimisation or gain-control optimisation — is limiting. The indication can be understood by referring to the lower and upper bound of the optimal tuning velocity (subsection 5.4.2). A high noise level, *i.e.*, much jitter, yields a less stable behaviour than a low noise level (Land, 1992). Consequently, high noise levels lead to a low upper bound of the optimal tuning velocity (c has a small value). Moreover, more jitter also leads to a lower average gain, due to attenuation (subsection 5.4.2). To maintain a target gain value, high jitter must be accompanied by a high lower bound of V_{opt}^* (c is large; cf. figure 5.14). In our definition of optimal tuning velocity in subsection 5.4.2 we set a hard lower bound (gain equals target gain) whereas the upper bound is softer (minimal response variance). Therefore, our results (increasing noise levels lead to increasing values of c) are explained by the fact that the lower bound of the optimal tuning velocity is increased when the noise level increases.

From the analysis of the results we derive that the noise level is one factor that determines the value of c and the limiting optimising factors. However, the stability of the system is determined by three factors: noise level, gain, and latency (Land, 1992). Increasing the target closed-loop gain G^* and the latency *Delay* also results in a less stable behaviour and, therefore, has similar effects on the lower and upper bound of the optimal tuning velocity when increasing the noise level. We conclude that the factors noise level, gain, and latency all set the optimal value of c and all determine what optimising factor is limiting — SNR optimisation or gain-control optimisation.

In summary, we investigated the third way in which the tuning velocity maximises the success of the gaze-stabilisation behaviour, *i.e.*, by optimising the SNR and the gain control of the model jointly.

5.4.4 Section conclusion

The aim of section 5.4 was to determine the optimal motion sensitivity that maximises gaze-stabilisation behaviour. Below we summarise how the optimal motion sensitivity was determined. Thereafter, we show that the determinant motion sensitivity causes a trade-off between spatial and temporal resolution in the model of gaze stabilisation.

The optimal motion sensitivity for gaze stabilisation

In the model, motion sensitivity is represented by the tuning velocity V^* of the augmented Reichardt correlator. To determine the optimal motion sensitivity we studied the three ways in which the optimal tuning velocity maximises the behavioural success of the model. First we studied how V^* optimises the SNR of the correlator output. Second we investigated how V^* optimises the gain control of the model. Third we studied the joint optimisation of SNR and gain control. To investigate the three ways we modelled open-loop and closed-loop control, created input-dependent and input-independent noise sources, and constructed an iterative update procedure to tune the closed-loop gain of the model.

We found that the optimal motion sensitivity (V_{opt}^*) is proportional to the retinal slip velocity (V_r) as represented by equation (5.18). In the equation, the value of the coefficient c depends on the extent to which each optimisation factor, *i.e.*, SNR and gain control, is limiting. We determined that the model's noise level, gain, and latency all determine which factor is limiting.

Motion sensitivity causes a trade-off

As a main finding of the section we show that motion sensitivity causes a trade-off between spatial and temporal resolution in the model of gaze-stabilisation. In section 5.3 it was established that the motion sensitivity is proportional to the fraction of temporal resolution over spatial resolution (equation (5.17)). In this section, we determined the relation between the optimal motion sensitivity and the velocity of image motion (equation (5.18)). Merging the equations (5.17) and (5.18) yields:

$$f_t^* = \frac{c}{a} \cdot (1 - G) \cdot V_b \cdot f_s^* \quad (5.24)$$

From equation (5.24) we derive that in our model of gaze stabilisation the optimal temporal resolution f_t^* is proportional to the optimal spatial resolution f_s^* . In figure 5.21 we draw four sample curves represented by equation (5.24). From the equation it follows that the optimal fraction of temporal resolution over spatial resolution, *i.e.*, the slope of each curve in the figure, depends on four variables: V_b , G , a , and c . For instance, a fast background velocity leads to a higher optimal fraction than slow background velocity, as is shown in the figure as well. A high fraction is obtained with high temporal resolution and low spatial resolution, whereas the inverse values result in a low fraction. Therefore, equation (5.24) suggests that the variables V_b , G , a , and c determine the position of the model's resolution on a trade-off curve.

Equation (5.24) optimises the motion sensitivity of the model of gaze stabilisation and describes a trade-off. Therefore, we conclude that the determinant motion sensitivity causes a trade-off between spatial and temporal resolution in the model of gaze-stabilisation. Important variables in the trade-off are the background velocity (V_b), the model gain (G), the coefficient a (determined by the band-pass filter widths; subsection 5.3.4), and the coefficient c which is determined by the noise level, gain,

Figure 5.21: The optimal temporal resolution f_t^* as a function of the optimal spatial resolution f_s^* for fast and slow background velocities.

and latency of the model (subsection 5.4.3). In the following sections, we study whether two other determinants, light sensitivity and contrast sensitivity, also cause a trade-off.

5.5 How visual resolution sets light sensitivity and contrast sensitivity

We show the relation between visual resolution, on the one hand, and the light sensitivity and contrast sensitivity, on the other. Establishing this relation is the first step in understanding how the determinants light sensitivity and contrast sensitivity cause a trade-off between spatial and temporal resolution (section 5.1.1).

The light sensitivity and contrast sensitivity are defined as the inverse of the minimal illumination and contrast of a scene needed for reliable detection (Wandell, 1995). In our model the reliability of detection is determined by the SNR which depends on the response amplitude of the augmented Reichardt correlator (section 5.4). Therefore, to determine the sensitivity as a function of resolution, we first measured the response amplitude as a function of resolution (subsection 5.5.1) and then as a function of illumination and contrast (subsection 5.5.2). Combining the results yields the light sensitivity and contrast sensitivity as a function of visual resolution (subsection 5.5.3). The section is summarised in subsection 5.5.4.

5.5.1 Response amplitude as a function of resolution

We measured the response amplitude of the augmented Reichardt correlator as a function of resolution. For this, we calculated the response amplitude as the maximal value R_m of the velocity-dependent response function: $R_m = \max(R(V)/S_0)$. To maintain a constant motion sensitivity, the response amplitude was measured as a function of resolution *scale*. The latter is explained below after which we show how the resolution scale sets the response amplitude.

Resolution scale

In section 5.4, it was shown that the pairs of resolution values that all optimise motion sensitivity lie on a straight line (figure 5.21). Motion sensitivity does not set the optimal scale of the resolution values, *i.e.*, the determinant does not determine the optimal pair that maximises the behavioural success. To determine the optimal pair, we kept the motion sensitivity fixed and measured the response amplitude R_m as a function of resolution *scale*. To obtain constant motion sensitivity we define the resolution scale S_R as follows:

$$S_R \equiv \frac{f_t^*}{f_t^N} = \frac{f_s^*}{f_s^N} \quad (5.25)$$

with the normalising resolution values f_t^N and f_s^N defined as:

$$f_t^N \cdot f_s^N \equiv 1 \quad (5.26)$$

The definition certifies that the fraction f_t^*/f_s^* equals the fraction f_t^N/f_s^N , *i.e.*, the motion sensitivity is constant, independently of S_R .

$$f_t^*/f_s^* = f_t^N/f_s^N \quad (5.27)$$

Using equations (5.25) and (5.26) the scale S_R is calculated as:

$$S_R = \sqrt{S_R^2} = \sqrt{\frac{f_t^*}{f_t^N} \cdot \frac{f_s^*}{f_s^N}} = \sqrt{f_t^* \cdot f_s^*} \quad (5.28)$$

Response amplitude as a function of resolution scale

We used equations (5.11) and (5.16) to calculate the response amplitude R_m as a function of the resolution scale S_R . The resolution values f_s^* and f_t^* were scaled by equation (5.25), $S_R \in [0.5, 2]$. We used the normalising values $f_t^N = 3.3$ Hz and $f_s^N = 0.3$ c/deg and fixed filter widths (see section 5.3). The scaled resolution values yielded scaled versions of the pre-filters G and H in the augmented Reichardt correlator. To obtain comparable results, we normalised the surface under the amplitude response functions of each scaled version of the filters to one. The results (figure 5.22) show that the response amplitude decreases when the resolution scale

Figure 5.22: The response amplitude R_m as a function of the resolution scale S_R .

increases. The relation between resolution scale S_R and response amplitude R_m was approximated as:

$$R_m = \frac{R_S}{S_R^2} \quad (5.29)$$

with the coefficient $R_S = 2.7 \cdot 10^{-7}$. Repeating the experiments for other values of the fraction f_t^N/f_s^N showed that the value of R_S is independent of the correlator's motion sensitivity.

The relation between resolution scale and response amplitude found here is explained as follows. Increasing the resolution scale results in increasing the values of spatial and temporal resolution, *i.e.*, attenuating the low-frequency components in the filtered image. According to the spatial power spectrum of natural images (5.15) the attenuation of low frequencies causes a reduction of the overall power in the filtered image and therefore a reduction of the response amplitude.

In summary, the response amplitude of the augmented Reichardt correlator depends on the resolution scale S_R in a non-linear fashion according to equation (5.29). The equation was obtained to understand how the visual resolution sets the light sensitivity and contrast sensitivity. As outlined above, the following subsection investigates how the illumination and contrast of natural images set the response amplitude.

5.5.2 Response amplitude as a function of illumination and contrast

To determine the response amplitude as a function of illumination and contrast we chose to calculate the factor S_0 in the spatial power spectrum of natural images (equation (5.15)) as a function of illumination and contrast. The motivation of the choice is follows. As was argued in section 5.2, the value of S_0 is determined by the mean illumination and average contrast in the image. Moreover, when rewriting equation (5.16) as $\overline{R(V)} = S_0 \cdot \int_0^\infty R_f \left(f_s, V \cdot f_s, \sqrt{1/f_s} \right) df_s$ we find that S_0 scales the velocity-dependent response of the augmented Reichardt correlator. Therefore, with a constant frequency-dependent response R_f , *i.e.*, constant resolution, the function that maps illumination and contrast to the response amplitude is determined by the factor S_0 .

To determine the factor S_0 we calculated the power spectrum of a set of 10 natural images introduced below. After that we show how the average contrast and mean illumination of the images are rescaled to calculate S_0 as a function of illumination and contrast.

Natural images

We used a set of 10 images obtained from Van Hateren's collection of natural stimuli³ (Van Hateren and Van der Schaaf, 1998). The image data, of which figure 5.23 shows an example, consisted of the raw light intensities (1536×1024 pixels) recorded by a carefully calibrated camera. The angle between two pixels in the pictures was approximately 1 minute of arc. We performed two operations on the dataset to make it suitable for our purposes. First, the image data were converted into a matrix of the original luminance values by multiplying the data by a calibration factor. This conversion sufficed for our measurements (*cf.* Van Hateren and Van der Schaaf, 1998). Second, we cropped the outer two lines of the image to avoid aberration problems.

To determine the value of S_0 , we calculated the mean spatial power spectrum using a 512-point fast Fourier transform (FFT). The spectra of 1020 horizontal lines in the image were subsequently averaged. Then we fitted equation (5.15) ($\alpha = 1$; 1-dimensional power spectrum) to the averaged power spectrum and calculated the optimal value of S_0 through least-squares analysis. Below, we show how the average contrast in the image was rescaled to calculate S_0 as a function of contrast.

Rescaling contrast

To calculate S_0 as a function of contrast, we rescaled the average contrast of the images in the set. The mean illumination was fixed at 1.0 cd/m^2 . There are various ways to measure the contrast of natural images (Peli, 1990; Van Hateren, 1993; Tolhurst and Tadmor, 1997). We employed the root-mean-square (RMS) measure

³See <http://hlab.phys.rug.nl/imlib/index.html>

Figure 5.23: An example of a natural image obtained from Van Hateren's image collection (Van Hateren and Van der Schaaf, 1998).

which is calculated as the standard deviation of the luminance values of all pixels in the image divided by the mean luminance value (Van der Schaaf and Van Hateren, 1996). The contrast was rescaled in three steps: (1) subtracting the mean luminance $\bar{I}$ from every pixel in the image; (2) multiplying every pixel value by a scaling factor $c \in [0.1, 9]$; (3) adding the original mean luminance $\bar{I}$. The new pixel values I' are thus calculated.

$$I'(x, y) = c \cdot (I(x, y) - \bar{I}) + \bar{I} \quad \text{with} \quad \bar{I} = \sum_x \sum_y I(x, y) / (X \cdot Y) \quad (5.30)$$

The constants X and Y represent the sizes of the cropped image. The contrast values obtained lay in a biologically relevant range (cf. Van der Schaaf and Van Hateren (1996)).

In figure 5.24 (a) we show the mean power spectra of the image of figure 5.23 for three different values of the RMS contrast. We find that the power increases with the RMS contrast as was expected (section 5.2). Figure 5.24 (b) depicts the calculated values of S_0 , averaged over 10 images, as a function of the RMS contrast. In the figures it is shown that increasing contrast values are matched to increasing values of S_0 in a non-linear fashion. The curve in figure (b) is approximated by the equation $S_0 = S_C \cdot C^2$ with $S_C = (2 \pm 1) \cdot 10^{-2}$ and C representing the RMS contrast.

We recall that the constant S_0 scales the velocity-dependent response of the augmented Reichardt correlator. Therefore, the response amplitude R_m is scaled as

Figure 5.24: The spatial power spectrum of the natural image in figure 5.23 for three values of image (RMS) contrast $\{0.1, 4.5, 9.0\}$, bottom-to-top order (a). The calculated values (mean and standard deviation) of the coefficient S_0 as a function of contrast (b).

a function of the RMS contrast as is shown in figure 5.24 (b). In the next section we determine the value of S_0 as a function of illumination.

Rescaling illumination

To calculate S_0 as a function of illumination we rescaled the mean intensity of the pixels in the image over a natural range of values, from a moonless, overcast night ($\bar{I} = 10^{-4}$ cd/m²) to clear day light ($\bar{I} = 10^4$ cd/m²) (Wandell, 1995). The RMS contrast is fixed at a base value 1. Of every scaled version of the image, we calculate the value of S_0 as described above. Figure 5.25 shows the calculated values of S_0 , averaged over the 10 images in the set, as a function of mean illumination. In the figure, it is shown that the value of S_0 changes non-linearly as a function of mean illumination. The curve is approximated by the equation $S_0 = S_1 \cdot \bar{I}^2$ with $S_1 = (2 \pm 1) \cdot 10^{-2}$ and $\bar{I}$ representing the mean illumination. We observe that the constant S_1 equals S_C , i.e., the response amplitude is scaled by illumination in the same way as by the RMS contrast. Below we explain the observation.

Rescaling illumination and contrast

We determined the factor S_0 when rescaling both the illumination and contrast of natural images. Above it was found that S_0 depends on C and $\bar{I}$ in the same way. This is explained as follows. The RMS contrast C is defined as the fraction of the standard deviation of illumination values σI divided by the mean illumination $\bar{I}$ ($C = \sigma I / \bar{I}$). When measuring S_0 as a function of contrast, we kept the mean illumination constant at value 1. Therefore, the RMS contrast equalled the standard

Figure 5.25: The calculated values (mean and standard deviation) of the coefficient S_0 as a function of mean illumination.

deviation ($C = \sigma I$). Similarly, when measuring S_0 as a function of illumination we used $C = 1$, *i.e.*, $\bar{I} = \sigma I$. Therefore, the response scale S_0 depends on the standard deviation of luminance values, *i.e.*, the product of the RMS contrast and the mean illumination ($\sigma I = C \cdot \bar{I}$), as follows.

$$S_0 = S_{CI} \cdot \sigma I^2 = S_{CI} \cdot (C \cdot \bar{I})^2 \quad (5.31)$$

Summary

We calculated the coefficient S_0 by rescaling the RMS contrast C and the mean illumination I of an image set. The results show that S_0 depends on C and I according to equation (5.31). Since S_0 scales the velocity-dependent response function, we have determined the response amplitude as a function of illumination and contrast.

5.5.3 Sensitivity as a function of resolution

We merge the results from subsections 5.5.1 and 5.5.2 to determine how the visual resolution sets the light sensitivity and contrast sensitivity of our model. To calculate the light sensitivity and contrast sensitivity, we aim at finding the minimal illumination and contrast required for reliable visual detection (cf. Wandell, 1995). In the beginning of this section, we showed that the reliability of visual detection depends on the response amplitude R_m of the augmented Reichardt correlator, *i.e.*, on the resolution scale S_R (subsection 5.5.1) and on the mean illumination and the

RMS contrast (subsection 5.5.2). Therefore, we calculated the threshold (minimal) illumination and contrast as a function of resolution scale.

The scaled response amplitude R_A of the correlator is defined as $R_A = S_0 \cdot R_m$. To calculate sensitivity, the equation is rewritten as follows. When substituting R_m and S_0 according to equation (5.29) and equation (5.31), we arrive at the following equation:

$$R_A = S_{CI} \cdot (C \cdot \bar{I})^2 \cdot \frac{R_S}{S_R^2} \quad (5.32)$$

We consider a threshold response amplitude θ_R that represents the minimal response needed for reliable detection. The expressions for light sensitivity and contrast sensitivity are obtained when substituting the threshold values of response amplitude ($R_A = \theta_R$) and either those of illumination ($\bar{I} = \theta_I$) or those of contrast ($C = \theta_C$) in equation (5.32).

$$\text{Light sensitivity} = \frac{1}{\theta_I} = \sqrt{\frac{R_S \cdot S_{CI}}{\theta_R}} \cdot \frac{C}{S_R} \quad (5.33)$$

$$\text{Contrast sensitivity} = \frac{1}{\theta_C} = \sqrt{\frac{R_S \cdot S_{CI}}{\theta_R}} \cdot \frac{\bar{I}}{S_R} \quad (5.34)$$

From the equations we derive that five factors determine light sensitivity and contrast sensitivity: (1) the response threshold θ_R ; (2) the RMS contrast C (in the case of light sensitivity); (3) the mean illumination $\bar{I}$ (in the case of contrast sensitivity); and the constants (4) R_S (subsection 5.5.1) and (5) S_{CI} (subsection 5.5.2).

In figure 5.26 we draw the light sensitivity and contrast sensitivity curves as a function of the resolution scale S_R . In the figure, the constants in equations (5.33) and (5.34) have the values calculated above, $R_S = 2.7 \cdot 10^{-7}$, $S_{CI} = 2 \cdot 10^{-2}$. θ_R was set to 0.01 to yield biologically plausible values of light sensitivity and contrast sensitivity. We calculated the light sensitivity using $C = 1$. The contrast sensitivity was calculated using $\bar{I} = 1 \cdot 10^3$ cd/m².

From our results we derive that the resolution scale S_R sets the light sensitivity and the contrast sensitivity of the augmented Reichardt correlator according to equations (5.33) and (5.34). The resolution scale can be calculated as the square root of the product of spatial and temporal resolution (equation (5.28)). Therefore, we conclude that the equations describe light sensitivity and contrast sensitivity as a function of the resolution.

5.5.4 Section summary

The aim of this section was to determine how the spatial and temporal resolution set the light sensitivity and the contrast sensitivity of the model of motion detection. To this aim, we measured the response amplitude of the augmented Reichardt correlator as a function of the resolution scale (the square root of the product of spatial and temporal resolution) and as a function of the illumination and contrast

Figure 5.26: The light sensitivity and contrast sensitivity of the model of gaze stabilisation as a function of the resolution scale S_R .

of natural images. After merging the results, we found that the light sensitivity and contrast sensitivity are inversely proportional to the resolution scale as described by equations (5.33) and (5.34).

The results help to understand how the determinants light sensitivity and contrast sensitivity cause a trade-off between spatial and temporal resolution in our model. In the following section we determine the optimal light sensitivity and contrast sensitivity, *i.e.*, the optimal product of resolution values.

5.6 The optimal light sensitivity and contrast sensitivity

By definition, the optimal sensitivity values are set by illumination and contrast. Therefore, we measure the optimal values of the determinants (*i.e.*, those values that maximise the behavioural success of our model) as a function of illumination and contrast. The optimal values may subsequently reveal whether the determinants cause a trade-off between spatial and temporal resolution.

In our model, illumination and contrast are represented by the value of S_0 (subsection 5.5.2). Moreover, the light sensitivity and contrast sensitivity are determined by the resolution scale S_R (subsection 5.5.3). Therefore, to find the optimal sensitivities we aim at finding the optimal value of S_R as a function of S_0 . To do so, we assume that S_R optimises the behaviour of gaze-stabilisation systems in three ways.

First, by maximising the SNR of the correlator output; second, by minimising the latency of the model; and third, by optimising the SNR and latency jointly. Comparably to the optimisation of motion sensitivity, we expect that the joint optimisation of the SNR and latency yields the optimal resolution scale, *i.e.*, the light sensitivity and contrast sensitivity that optimise gaze stabilisation.

In subsection 5.6.1, we test our assumption that S_R optimises the behaviour of gaze-stabilisation systems by maximising the SNR and minimising the latency. Subsequently, we determine the optimal S_R as a function of S_0 in subsection 5.6.2. Subsection 5.6.3 summarises our findings.

5.6.1 Optimisation of SNR and latency

Below, it is argued that the resolution scale S_R optimises the behaviour of gaze-stabilisation systems by maximising the SNR and minimising the latency. First we discuss SNR maximisation, followed by latency minimisation.

SNR maximisation

In chapter 4, we showed that the resolution scale optimises the behaviour of biological and artificial position-detection systems by maximising the SNR in the system. We presume that the resolution scale optimises the behaviour of gaze-stabilisation systems in a similar way. The resolution scale sets the response amplitude R_m (the maximal velocity-dependent response) of the augmented Reichardt correlator as described by equation (5.29). The SNR of the noisy output signal is determined by the noise level M and the velocity-dependent response R (equation (5.23)). Hence, the resolution scale sets the SNR of the noisy response signal R' . Moreover, the SNR of the noisy response signal R' determines the response variance of the model, *i.e.*, the quality of the behaviour. According to equation (5.29), the response amplitude (and therefore the SNR) is optimised when the resolution scale is minimised (see also figure 5.22). Hence, we conclude that the minimal resolution scale S_R optimises the behaviour of gaze-stabilisation systems by maximising the SNR.

Latency minimisation

The latency of the model of gaze stabilisation is minimised when the delay block in the control diagram (figure 5.1) is minimised and when the time constants of the low-pass temporal filters (the pre-filters in the augmented Reichardt correlator and the integrating filter LP) are minimised, *i.e.*, when the temporal resolution is maximised. The resolution scale is calculated as the square root of the product of spatial and temporal resolution (equation (5.28)). Therefore, assuming that the contribution of the filters' time constants to the model's latency is comparable or larger than the contribution of other delay terms, we may conclude that the latency is optimised when the resolution scale is maximised.

To provide evidence for the assumption we remark that the latency of biological gaze-stabilisation systems depends on the illumination (Maunsell *et al.*, 1999) and

the contrast (Warzecha and Egelhaaf, 2000) of moving natural images. Warzecha and Egelhaaf believe cellular mechanisms to cause the changes in latency, *i.e.*, they suppose adaptation to the illumination and contrast level. For instance, increased contrast and illumination give rise to a reduced input resistance of neurons in motion-detection pathways. The reduction of input resistance results in a lowered response latency. Moreover, since the time constant τ is calculated as the product of the resistance and capacitance of an RC network (equation (5.4)) the reduction also leads to a higher tuning frequency of the filter, *i.e.*, higher temporal resolution and higher resolution scale S_R . The range of illumination and contrast levels used by Warzecha and Egelhaaf (2000) gives rise to changes in response latency of the complete visuo-motor system over a significant range of one order of magnitude. We may conclude that the changes in latency are induced by changes in the resolution scale.

Altogether we found that the behaviour of gaze-stabilisation systems is optimised when the SNR is maximised and the latency is minimised. The SNR is maximised when the resolution scale is *minimised* while the model's latency is minimised when the resolution scale is *maximised*.

5.6.2 The optimal resolution scale

We determine the resolution scale S_R^* that optimises gaze stabilisation. In subsection 5.6.1 we showed that S_R optimises the SNR and latency of the system. Below, we discuss the implementation of the SNR and the latency as a function of resolution scale. Thereafter, we describe the experimental set-up that was used to determine S_R^* . We provide our expectations before presenting the results.

SNR and latency

We implemented the SNR and latency as a function of resolution scale S_R as follows. In congruence to section 5.4 we implemented the noise as a Poisson process. The process was modelled by drawing variables from a Gaussian distribution with mean R and standard deviation $\sqrt{R}$ (equation (5.23)). The SNR of the noisy signal R' equalled $\sqrt{R/M}$, *i.e.*, the SNR increased when the signal R increased or when the noise level M decreased. The response of the augmented Reichardt correlator, *i.e.*, the signal R , was determined by the resolution scale S_R (subsection 5.5.1). The SNR is thus implemented as a function of resolution scale.

The latency of the model depends on by the term *delay* and the temporal transient aspects of the pre-filters in the correlator and of the low-pass filter LP (see figure 5.1)). Since we simulated only the steady-state response of the augmented Reichardt correlator (subsection 5.4.2), we implemented all the temporal transient aspects of the model in the low-pass filter LP (see the control diagram in figure 5.1). To do so, we assigned a different value to the time constant τ_3 of the filter LP , for each value of S_R used. The constant τ_3 was assigned the sum of a base value T and the time constant τ_1 of the pre-filter H in the augmented Reichardt correlator. The

value of τ_1 determines the temporal resolution of the model and, therefore, the scale S_R . In this way, we implemented the latency as a function the resolution scale S_R . The value of T was set such that the resolution scale used in section 5.4 corresponded to the same value τ_3 as was used in that section.

Experimental set-up

To determine the optimal resolution scale, the behaviour of the model was measured for a range of values of S_R and S_0 . We determined the optimal value of S_R for every value of S_0 used. The optimal resolution scale S_R^* was defined as the resolution scale that yielded the target closed-loop gain G^* and produced minimal response variance (*i.e.*, the least jitter).

Comparably to section 5.4 we calculated the response of the augmented Reichardt correlator using a lookup table containing the correlator gain values K_c for a range of values of the input velocity V_r and the resolution scale S_R . The motion sensitivity thus remained constant while the light sensitivity and contrast sensitivity varied as a function of S_R . In all other respects, the experimental set-up reflected the set-up used in subsection 5.4.3.

Expected results

Using this set-up we expected the following results. In section 5.4, we set the factor S_0 and the response amplitude R_m to 1. Therefore, when optimising motion sensitivity in section 5.4 the scaled response amplitude equalled $R_A = S_0 \cdot R_m = 1$. In conformity, we expected that, also in this section, the optimal resolution scale sets the response amplitude R_m such that $R_A = 1$, *i.e.*, $R_m = R_A/S_0 = 1/S_0$. Therefore, using equation (5.29), we expected the following relation between the optimal resolution scale S_R^* and the factor S_0 :

$$S_R^* = \sqrt{R_S \cdot S_0} \quad (5.35)$$

Below we present the results and compare them to the expected results.

Results

Using S_0 values ranging from $2 \cdot 10^{-10}$ ($\bar{I} = 10^{-4}$ cd/m²; dark night) to $2 \cdot 10^6$ ($\bar{I} = 10^4$ cd/m²; bright daylight), we obtained the optimal values S_R^* shown in figure 5.27. We fitted the dashed curve to the data by least-squares analysis. The fitted curve is described by equation (5.36).

$$S_R^* = b \cdot S_0^\alpha \quad (5.36)$$

with $b = 2.7 \cdot 10^{-6}$ and $\alpha = 0.16$.

As our main result we found that the optimal resolution scale, *i.e.*, the optimal light sensitivity and contrast sensitivity, is determined by the factor S_0 , *i.e.*, the illumination and contrast of natural images.

Figure 5.27: The optimal resolution scale S_R^* as a function of S_0 .

When comparing the fitted curve to the anticipated curve we now determine that the results differ from the expected outcomes in two main ways. First, the optimal values of the resolution scale are much smaller than expected since the factor $b = 2.7 \cdot 10^{-6}$ (equation (5.36)) is smaller than the anticipated value $\sqrt{R_S} = 5.2 \cdot 10^{-4}$ (equation (5.35)). Earlier in this subsection, we showed that the expected resolution scale produced a scaled response amplitude $R_A = 1$. The smaller resolution scales that were measured yielded higher response amplitudes and, therefore, higher SNRs, but also higher response latencies (subsection 5.6.1). Therefore, the results show that the behaviour of the model is optimal when the scaled response amplitude R_A is larger than 1, *i.e.*, when the SNR is better and the latency is worse than in section 5.4.

Second, equation (5.27) is not a square-root function, since we found an exponent $\alpha = 0.16$ (equation (5.36)). This relation indicates that for low values of S_0 the optimal resolution scale is higher than expected, whereas for high S_0 values the optimal resolution scale is lower than expected. As was shown in subsection 5.6.1, a high resolution scale results in low latency and low SNR. Therefore, the results show that for low values of S_0 the system optimises the latency more than the SNR (*i.e.*, latency optimisation is limiting), whereas for high values of S_0 the system optimises the SNR more than latency (*i.e.*, SNR optimisation is limiting).

Summary

We determined the resolution scale S_R^* that optimises the behaviour of the model of gaze stabilisation. To determine S_R^* we first implemented the SNR and latency as a function of S_R . Thereafter, we explained the experimental set-up to optimise the behaviour and showed what results were expected. We found that the optimal resolution scale S_R^* depends on the factor S_0 according to equation (5.36). We also found that the behaviour of our model is optimal when $R_A > 1$, *i.e.*, when the SNR is better and the latency is worse than in section 5.4. Finally, it was derived from the results that the factor S_0 determines which optimisation factor (SNR maximisation or latency minimisation) is limiting.

5.6.3 Section summary

The aim of this section was to determine the optimal light sensitivity and contrast sensitivity of the model of gaze stabilisation. Below we summarise how the optimal sensitivity values were determined and how they may be calculated as a function of the illumination and contrast of natural images. Thereafter, we show that the determinants light sensitivity and contrast sensitivity cause a trade-off between the spatial and temporal resolution in our model of gaze stabilisation.

The optimal light and contrast sensitivity

In our model the light sensitivity and contrast sensitivity are represented by the resolution scale S_R (see equations (5.33) and (5.34)). To determine the optimal sensitivity values we tested (subsection 5.6.1) our assumption that the resolution scale optimises the behaviour of gaze stabilisation in three ways: by maximising the SNR; by minimising the latency; and by optimising the SNR and latency jointly. To find the optimal resolution scale S_R^* , we studied the joint optimisation of SNR and latency in subsection 5.6.2. The value of S_R^* was determined as a function of the factor S_0 described by equation (5.36). The factor S_0 represents the illumination and contrast in natural images (equation (5.31)). Therefore, equation (5.36) determines the optimal resolution scale as a function of illumination and contrast of natural images. It was also derived from the results that the optimal values of SNR and latency are larger than the values used in section 5.4. Finally, we found that the value of S_0 determines which optimisation factor (SNR or latency) is limiting.

Light sensitivity and contrast sensitivity cause a trade-off

Below we show that the determinants light sensitivity and contrast sensitivity cause a trade-off between the spatial and temporal resolution of our model of gaze stabilisation. Light and contrast sensitivity are represented by the resolution scale S_R , which was defined as the square root of the product of spatial and temporal resolution (equation (5.28)). In section 5.6, we determined the relation between the optimal resolution scale S_R^* and the factor S_0 (equation (5.36)). Equation (5.31)

shows that S_0 depends on illumination $\bar{I}$ and contrast C of natural scenes. Merging equations (5.28), (5.31), and (5.36) yields the optimal values of spatial and temporal resolution as a function of illumination and contrast:

$$f_t^* = \frac{b^2 \cdot S_{CI}^{2\alpha} \cdot (C \cdot \bar{I})^{4\alpha}}{f_s^*} \quad (5.37)$$

Equation (5.37) indicates that the optimal temporal resolution f_t^* is inversely proportional to the spatial resolution f_s^* . In figure 5.28 we draw four sample curves described by the equation. The curves in the figure describe combinations of spa-

Figure 5.28: The temporal resolution f_t^* as a function of spatial resolution f_s^* that optimise light sensitivity and contrast sensitivity.

tial and temporal resolution that optimise the light and contrast sensitivity of the model of gaze stabilisation. Moreover, the curves resemble the trade-off curves that were shown in chapters 3 and 4. Therefore, we may conclude that the spatial and temporal resolution of our model lie on a trade-off curve to optimise the light sensitivity and contrast sensitivity of the model and thus maximise its behavioural success. According to equation (5.37), the trade-off curves are scaled by the product $b^2 \cdot S_{CI}^{2\alpha} \cdot (C \cdot \bar{I})^{4\alpha}$. In this product the factors b and S_{CI} are constants. Therefore, the illumination and contrast scale the trade-off curves caused by light sensitivity and contrast sensitivity.

5.7 A trade-off in motion-detection systems

The overall aim of this chapter is to understand whether the determinants motion sensitivity, light sensitivity, and contrast sensitivity cause a trade-off between spatial and temporal resolution in motion-detection systems. To achieve the aim, we constructed a model of gaze stabilisation (section 5.2) and studied how the determinants set the optimal resolution values of our model (sections 5.3 to 5.6). In this section we merge the experimental results to understand how the determinants cause a trade-off.

In subsection 5.4.4 we concluded that the fraction of temporal resolution over spatial resolution is set by the typical velocity of moving images to optimise the motion sensitivity of the model as described by equation (5.24). The equation represents straight lines of resolution pairs as was shown in figure 5.21. Later, in subsection 5.6.3, we concluded that illumination and contrast of natural images set the product of spatial and temporal resolution that optimises the light sensitivity and contrast sensitivity of the model as described by equation (5.37). The equation represents curved lines of resolution pairs as was shown in figure 5.28.

We redraw the lines from figures 5.21 and 5.28 in figure 5.29. From the results we infer that the optimal resolution values lie at the intersections of the lines in this figure. The locations of the intersections are calculated by merging the equations (5.24)

Figure 5.29: Temporal resolution as a function of spatial resolution that optimise the motion sensitivity, light sensitivity, and contrast sensitivity of the model of gaze stabilisation.

and (5.37) into the following equations:

$$f_s^* = b \cdot S_{CI}^a \cdot \sqrt{\frac{a}{c \cdot (1 - G)}} \cdot \frac{(C \cdot \bar{I})^{2\alpha}}{\sqrt{V_b}} \quad (5.38)$$

$$f_t^* = b \cdot S_{CI}^a \cdot \sqrt{\frac{c \cdot (1 - G)}{a}} \cdot \sqrt{V_b} \cdot (C \cdot \bar{I})^{2\alpha} \quad (5.39)$$

Equations (5.38) and (5.39) describe the optimal values of spatial and temporal resolution, respectively, as a function of typical image velocity V_b , illumination $\bar{I}$, contrast C , and five other factors as explained below.

- The factor b is the coefficient in the relation between S_R^* and S_0 (equation (5.36)). Its value is set by the optimal combination of SNR and latency of the model.
- The factor S_{CI} is the coefficient in the relation between S_0 and the product $(C \cdot I)$ (equation (5.31)). Its value was established at $2 \cdot 10^{-2}$.
- The factor a is the coefficient in the relation between the fraction of temporal resolution over spatial resolution and the tuning velocity (equation (5.17)). Its value is greater than 1, depending on the pass-band widths of the spatial and temporal pre-filters.
- The factor c is the coefficient in the relation between the optimal tuning velocity and the image velocity (equation (5.18)). Its value depends on the noise level in the system, as well as on the gain and latency.
- The factor G is the closed-loop gain of the system. It depends on the open-loop gain g (equation (5.13)) and the level of gain control in the system (subsection 5.4.2).

Based on equations 5.38 and 5.39 we may conclude that the determinants motion sensitivity, light sensitivity and contrast sensitivity cause a trade-off between spatial and temporal resolution in a motion-detection system (our model of gaze stabilisation). The optimal resolution values lie on a trade-off curve that is scaled by illumination and contrast of natural images. The optimal position is determined by the velocity of the moving images. We found five other factors that co-determine the optimal resolution values.

5.8 Chapter discussion and conclusions

Below, we discuss our results and present our main conclusions. In the previous section, the results of our experiments were epitomised. They demonstrate (1) that the velocity of moving images sets the optimal visual resolution on a trade-off curve and (2) that illumination and contrast of natural images set the scaling factor of the curve. From our observations, we may conclude that the results presented in the

current chapter generalise the findings of chapter 4 to the case of motion-detection systems.

In subsection 4.8.1 we argued that the results are applicable to biological position-detecting systems. We deal with the current results in analogy. In section 5.8.1, we compare our model output to the response of biological systems. From the comparison we may determine to what extent our results predict the trade-off in biological motion-detection systems. In subsection 5.8.2 we compare our model to other models of motion detection to show our model's strengths and weaknesses. In subsection 5.8.3 we present our conclusions.

5.8.1 Comparison to biological systems

To demonstrate the applicability to real animals, the response of our model is compared to that of biological systems. In section 5.2 we showed that our model captures the basic characteristics of biological motion detectors. Therefore, the results obtained are expected to apply to biological systems as well. To support the validity of the model even better, we compared our model to biological motion-detecting systems both quantitatively and qualitatively. Below we first compare the response of the model and of animals using biologically plausible parameter settings. Thereafter, we compare the qualitative shape of the response curves. Finally, we show to what extent the results obtained with the model apply to biological systems.

Quantitative comparison

The tuning frequencies of the model and of biological motion-detection systems are quantitatively matched. This is shown as follows. The frequency-dependent response surface in figure 5.3 is described by equation (5.11). The parameter values used ($\Delta\varrho_1 = 0.9$ deg, $\Delta\varrho_2 = 4.5$ deg, $\Delta\phi = 1.4$ deg, $\tau'_1 = 0.05$ s, and $\tau'_2 = 0.1$ s) were chosen to lie in a biologically plausible range. For instance, the angular receptor distance of the blowfly *Calliphora* equals $\Delta\phi = 2.4$ deg, whereas for human photoreceptors in the fovea holds $\Delta\phi = 0.007$ deg (Land, 1981); the typical value of $\Delta\varrho_1$ in flies is about 1.5 deg (Hardie, 1985) but it is much lower in the human eye due to the low angular receptor distance; the cut-off frequency of the light-adapted fly photoreceptor response is approximately 20 Hz (Juusola and Hardie, 2001), which corresponds to $\tau'_1 = 0.014$ s; however, τ'_1 can reach values over 0.3 s in vertebrates and in dark-adapted eyes (cf. Warrant, 1999). In figure 5.3, the tuning frequencies are $f_s^* = 0.20$ c/deg and $f_t^* = 2.3$ Hz. These values are comparable to the tuning frequencies of *Calliphora* that are about $f_s^* = 0.13$ c/deg and $f_t^* = 5$ Hz (Mastebroek *et al.*, 1980). The tuning frequencies of the model are, therefore, quantitatively comparable to the tuning frequencies of biological systems.

Qualitative comparison

Qualitatively, we showed in section 5.2.3 that the shape of the frequency-dependent response surface (figure 5.3) corresponds to the response measured in flies. Unfortu-

nately, the spatio-temporal frequency response of vertebrate motion-detecting neurons is not available. Yet, the contrast-sensitivity graphs measured in the early stages of the visual systems of cats (Holub and Morton-Gibson, 1981; Crawford *et al.*, 1990), humans (Anderson and Burr, 1985; Crawford *et al.*, 1990), and flies (Van Hateren, 1992b) display pass-band characteristics around the tuning frequencies, which corresponds to our model's response. However, the contrast-sensitivity graphs are not space-time separable. For instance, further away from the tuning frequencies, the graph has a low-pass characteristic. This phenomenon is not demonstrated by our model, which is probably due to the simple types of filters we use (cf. Van Hateren, 1992b; Van Hateren, 1993; Dror *et al.*, 2001). The use of more realistic filters seems an interesting topic for further study.

Applicability to biological systems

To show the applicability of our results to biological systems we refer to the observations above. There we showed that the frequency-dependent response function (equation (5.11)) of the augmented Reichardt correlator matches the characteristics of biological motion detectors both quantitatively and qualitatively. The results obtained with our model depend on the frequency-dependent response function as well as on the response of other components that were previously shown to match their biological sources (section 5.2). Therefore, we presume that the results apply to biological motion-detection systems. Moreover, above it was shown that the results compare to both invertebrate and vertebrate animal species. Therefore, we conclude that the model of gaze stabilisation is a good model for studying the trade-off between spatial and temporal resolution in biological motion-detecting systems.

5.8.2 Comparison to other models

To show the strengths and weaknesses of our particular model we compare it to other models of biological motion detection. It is shown that our model distinguishes itself from other models in two ways: (1) our model employs band-pass filters; (2) we incorporate the symmetric type of Reichardt correlator in our model. Both differences are discussed below, followed by a discussion of our model's strengths and weaknesses.

Band-pass filters

In figure 5.2, G and H denote spatial and temporal band-pass filters of the augmented Reichardt correlator, respectively. As was mentioned before, our correlator is based on other correlator models, many of which only use low-pass pre-filters (O'Carroll *et al.*, 1997), or band-pass filtering in just the temporal domain (Kern *et al.*, 2001a; Neumann and Bülthoff, 2001). In contrast, the augmented correlator comprises both spatial and temporal band-pass filters (cf. Van Hateren, 1992b; Van Hateren, 1993; Snippe and Koenderink, 1994). The use of band-pass

filters indicates that the optimal resolution is determined by more factors than just image velocity. Namely, in subsection 5.3.4 we have shown how the pass-band widths of the spatial and temporal filters determines the value of the coefficient a in equation (5.17). Equations (5.33) and (5.34) show that the coefficient a partly determines the optimal resolution values. As a result, we found that the band-pass widths co-determine the optimal resolution. This finding indicates that our model is more general than other models that only use low-pass filtering. The results show how animals might use the pass-band width as an additional mechanism to adjust their optimal resolution. Additional mechanisms are important, since animals need their visual system for various behaviours. Different behaviours require different resolutions, *e.g.*, face recognition requires high spatial resolution. The additional mechanisms offer a means to determine the optimal eye design for each type of visually guided behaviour.

Symmetric and asymmetric correlators

Besides by the use of band-pass filters, the augmented Reichardt correlator differs from other models in a second respect, being the use of symmetric input arms. Although most studies agree that the symmetric correlator describes the output of biological motion detectors well, Srinivasan, Poteser, and Kral (1999) assumed that animals use asymmetric correlators (see figure 2.2.2) for visually guided behaviours such as depth estimation, flying through narrow gaps, and the regulation of flight velocity. The reason for assuming asymmetric rather than symmetric correlators is the fact that the former are true estimators of velocity, independent of the spatial pattern of the environment. In contrast, our symmetric correlator is selective for temporal frequency, rather than velocity (see equation (5.11) and Dror *et al.*, 2001). Identified neurons in the fly brain that are thought to be responsible for the control of gaze-stabilisation behaviour (Hausen, 1982; Srinivasan *et al.*, 1999) are frequency-dependent: their output is best described by symmetric correlator models. Therefore, the use of symmetric models is preferred when studying gaze-stabilisation behaviour. In further studies on other types of visually guided behaviour, we expect to find a similar relation between image velocity and the resolution of asymmetric correlator models as was found in the current chapter (see also chapter 7).

Strengths and weaknesses

From the comparisons in subsections 5.8.1 and 5.8.2 we derive three points of strengths and weaknesses of our model of gaze stabilisation. First, the results obtained with our model are argued to apply to biological systems. Therefore, although not all details of biological motion detection are captured, our model is a good representation of gaze-stabilisation behaviour. Second, the use of band-pass filters provides additional insight into the optimal eye design. Third, although the model's symmetric Reichardt correlators represent the mechanisms of gaze stabilisation well, probably not all types of visually guided behaviour are captured by the

model. However, we expect that our results can be generalised to other behaviours as well (see chapter 7).

5.8.3 Chapter conclusions

The aim of the chapter was to determine whether the determinants motion sensitivity, light sensitivity, and contrast sensitivity cause a trade-off between spatial and temporal resolution in motion-detection systems. To investigate this, we constructed a model of a common type of visually guided behaviour controlled by a motion-detection system: the model of gaze stabilisation (section 5.2). Subsequently, we determined how the determinants motion sensitivity (sections 5.3 and 5.4) and the determinants light sensitivity and contrast sensitivity (sections 5.5 and 5.6) set the optimal resolution of our model, *i.e.*, the resolution that maximised the behavioural success of the model. In section 5.7 we merged the results to show that the determinants cause a trade-off between spatial and temporal resolution in motion-detection systems. Finally, in section 5.8 we generalised our findings to biological motion-detection systems and to the findings from chapter 4 on position-detection systems. Below we present our main conclusions.

Conclusions

From the results in this chapter we are able to draw a series of eight conclusions.

1. The determinants motion sensitivity, light sensitivity, and contrast sensitivity cause a trade-off between the spatial and temporal resolution of motion-detection systems. In addition, the two remaining determinants, heat production and energy consumption, are expected to cause a trade-off as well (section 5.1).
2. The optimal resolution values are determined by the velocity, illumination, and contrast of moving natural images as well as by five other factors.
3. The motion sensitivity optimises the gaze-stabilisation behaviour by maximising the SNR and gain control of our model. The limiting optimisation factor is determined by the level of response variance.
4. The light sensitivity and contrast sensitivity optimise the visually guided behaviour by maximising the SNR and minimising the latency of the model of gaze stabilisation. The limiting optimisation factor is determined by the illumination and contrast of natural images.
5. The motion sensitivity is determined by the resolution and band-pass filter width of the model.
6. The light and contrast sensitivity are determined by the resolution scale, the illumination (for contrast sensitivity), the contrast (for light sensitivity), and the response threshold.

7. The model of gaze stabilisation is an appropriate model to determine the optimal resolution of biological motion-detection systems.
8. We expect that all conclusions mentioned apply to both artificial and biological motion-detection systems.

In summary, in this chapter we have shown that the determinants that were identified in section 3.2 cause a trade-off between spatial and temporal resolution in motion-detection systems. The results generalise the findings of chapter 4 where we investigated the determinants that cause a trade-off in position-detection systems.

Chapter 6

The importance of being situated

“This matter may prove to be one of vital importance.”

*Lady Bracknell in ‘The importance of being Earnest’
by Oscar Wilde (1854 – 1900), Anglo-Irish playwright.*

Chapter contents: Comparison to controlled behavioural studies, comparison to information theory, the optimality of situated modelling

In chapter 2 we selected situated systems modelling as the approach for our research. This chapter revisits the appropriateness of our approach. In particular, we argue that our experimental results (chapters 4 and 5) show that the situated approach is more appropriate for studying visually guided behaviour than the two main alternative approaches, *i.e.*, controlled behavioural studies and information theory (section 2.5).

Section 6.1 provides a qualitative comparison of our results with those of controlled behavioural animal studies. In section 6.2 we present a quantitative comparison of the application of situated modelling and information theory. The chapter is concluded in section 6.3 by stating that our approach is more appropriate for studying visually-guided behaviour than controlled behavioural studies and information theory.

6.1 Comparison to controlled behavioural studies

According to our discussion in section 2.5, the class of controlled behavioural studies represents one of the two main alternatives to our research approach (see also

figure 2.1). We compare our results obtained with situated modelling qualitatively to those obtained with controlled behavioural studies. The comparison is performed in terms of the three topics discussed in section 2.3 concerning animal research in general: (1) comparability of the results (subsection 6.1.1), (2) the availability of relevant data (subsection 6.1.2), and (3) the provision of insight into the causes of animal behaviour (subsection 6.1.3).

6.1.1 Comparability of the results

Our models of target tracking (section 4.2) and gaze stabilisation (section 5.2) each employ a single measure of visual resolution, thus facilitating the study of the trade-off. In contrast, animal research employs a wide variety of measures of visual resolution (*e.g.*, flicker fusion frequency, spatial and temporal cut-off frequencies; Laughlin, 1981) which are often difficult to relate for studying the trade-off (see also chapter 7).

6.1.2 Availability of relevant data

Our situated models of visually guided behaviour do not impose limitations on the availability of relevant experimental data. Simulation studies offer an unlimited source of data with computational resources as the only cost. Therefore, the results of modelling studies offer statistical reliability.

In contrast, animal studies are characterised by a limited availability of data due to the time-consuming procedures involved in obtaining these data from real animals. As a case in point, O'Carroll *et al.* (1996) published the spatial and temporal resolution data of three insect species. The data showed that the motion sensitivities of these animals are matched to their visual ecology. However, due to the small size of the data-set, a statistically sound analysis of the findings and a successful generalisation to other biological visual systems is unfeasible.

6.1.3 Provision of insight

Our situated approach enables a fully transparent model that allows for a detailed study of the behaviour and its underlying processes. Hence, the factors causing behaviour can be identified. This is much more difficult, if not impossible, in animal studies. The aforementioned work by O'Carroll *et al.* (1996) provides an illustrative example. Their animal approach yields correlations between visual resolution and visual ecology which suggest that a resolution trade-off exists. However, the study does not explain which are the actual causes that lead to the observed relation between visual resolution and visual ecology.

In summary, our qualitative comparison shows that the three disadvantages of animal research are resolved by our modelling approach. Therefore, we argue that situated systems modelling is more appropriate than controlled behavioural studies to investigate the trade-off between spatial and temporal resolution in visually guided behaviour.

6.2 Comparison to information theory

In section 2.5 we identified information theory as the other main alternative to situated systems modelling (see also subsection 2.4.1). Below, we argue that our approach offers a better insight into visually guided behaviour than the information-theoretical approach by comparing both approaches quantitatively.

In section 5.4 we optimised the motion sensitivity of the augmented Reichardt correlator using situated modelling. Alternatively, the sensitivity of the correlator can be optimised using information theory. The latter approach views the correlator as an information channel of which the optimal sensitivity is found by maximising its information capacity (subsection 2.4.1). In subsection 6.2.1 we show how this is done, followed by the presentation of our results in subsection 6.2.2. Subsection 6.2.3 draws conclusions from our quantitative comparison.

6.2.1 Maximising information capacity

The information capacity of an information channel is maximised by maximising the fraction $S(\nu)/N(\nu)$, *i.e.*, the SNR, of the channel's output signal (see equation (2.1)). Therefore, when using information theory, the optimal motion sensitivity of the augmented Reichardt correlator is obtained by maximising the SNR of the correlator's output signal. Since besides the augmented Reichardt correlator, all components in our model (figure 5.1) are linear, we approximate the SNR as the fraction of the mean retinal slip velocity over the standard deviation of the slip velocity: $\text{SNR} = \bar{V}_r / \sigma V_r$. To optimise the motion sensitivity, we calculate the information capacity for a range of values of the tuning velocity and ten background velocities. The tuning velocity that maximises the information capacity determines the optimal motion sensitivity.

6.2.2 Results

The lower curve in figure 6.1 shows the optimal tuning velocity of the augmented Reichardt correlator as a function of background velocity for the information-theoretical approach. The upper curve depicts the optimal tuning velocity for the situated approach and is reproduced from subsection 5.4.2 (figure 5.19).

The results clearly show that the values yielded by information theory differ from the values obtained by situated modelling. For all background velocities examined, the optimal tuning velocity is an order of a magnitude lower for the information-theoretical approach than for the situated approach. Our situated approach optimised the tuning velocity by maximising the behavioural success of the model of gaze stabilisation (subsection 5.1.1). As a consequence, our results show that maximising the information capacity of the augmented Reichardt correlator does not lead to the maximisation of behavioural success.

The different results obtained by both approaches can be readily explained in terms of the different targets of optimisation. The information-theoretical approach optimises the SNR to maximise the information capacity, whereas situated approach

Figure 6.1: The optimal tuning velocity of the model of gaze stabilisation as a function of background velocity. The upper curve shows the results obtained with situated modelling. The lower curve was determined by employing information theory.

optimises the closed-loop gain to maximise the success of gaze-stabilisation behaviour. Figure 6.2 (a) shows the information capacity for both approaches averaged over the ten background velocities (*i.e.*, the data points in figure 6.1). Evidently, the information-theoretical approach yields a larger information capacity than the situated approach.

However, as shown in figure 6.2 (b) the situated approach is associated with the largest closed-loop gain. Notably, the low optimal tuning velocities, obtained by information theory, yield more gain control than the higher tuning velocities obtained by situated modelling (see section 5.4.2). As a result, the approach of information theory leads to an excess gain control and a reduced closed-loop gain. In contrast, the situated approach does not yield this problem.

6.2.3 Conclusions

In summary, we state that whereas the situated approach performs poor in terms of information capacity, it does lead to better performance in terms of behaviour.

Clearly, from an evolutionary perspective, behavioural performance is of primary importance for survival. In chapter 1 we assumed that the trade-off was caused by the evolutionary adaptations to the visual ecology of animals. Therefore, the quantitative comparison of situated modelling and information theory reveals that the latter is of limited value for studying the trade-off. Optimal information capacity

Figure 6.2: (a) The mean and standard deviation of the information capacity of the Reichardt correlator obtained with information theory and situated modelling. (b) The mean and standard deviation of the closed-loop gain obtained with information theory and situated modelling. The average values were calculated using the optimal values of tuning velocity shown in figure 6.1.

does not correspond to optimal behavioural performance. Hence we conclude that the situated approach is more appropriate for studying visually guided behaviour.

6.3 The importance of situated systems modelling

In this chapter, our results obtained with the situated approach were compared to those of the two main alternative approaches: controlled behavioural studies and information theory. We argued that our modelling results are better comparable, are available in larger quantities, and provide a better insight into the resolution values than the results of animal research do (section 6.1). Moreover, our situated approach leads to results that better conform to the principles of evolutionary adaptation than information theory does (section 6.2). Based on the comparisons, we may conclude that situatedness is vital to studying the resolution of visually guided behaviour.

Chapter 7

Generic causes of the trade-off

A good cause can become bad if we fight for it with means that are indiscriminately murderous.

Freeman Dyson (1923 – now), British-born U.S. physicist, author.

Chapter contents: Determinants of the trade-off, additional determinants, occurrences of visually guided behaviour, application to biological systems, application to artificial systems, conclusions and future research

To draw conclusions from our research we aim at working towards an exhaustive list of generic causes of the trade-off between spatial and temporal resolution. In chapter 1 we set three criteria for generic causes. They (1) explain the existence of a trade-off in many occurrences of visually guided behaviour. Moreover, they apply to many instances of (2) biological visual systems and (3) artificial visual systems.

To work towards the generic causes we start in section 7.1 by iterating our five determinants and recalling how they cause a trade-off in the selected models of visually guided behaviour (chapters 4 and 5). From our research we derived some additional determinants which are summarised in section 7.2. Then, it is verified whether the determinants meet the three criteria of generic causes. First, in section 7.3 we discuss alternative occurrences of visually guided behaviour. Second, the determinants are applied to a range of biological visual systems (section 7.4). Third, section 7.5 does the same for artificial visual systems. Section 7.6 presents the generic causes of the trade-off.

7.1 The determinants iterated

In chapter 3 we established five determinants of the trade-off between spatial and temporal resolution. The determinants were optimised to accomplish maximal behavioural success in a model of a position-detection system (chapter 4) and of a motion-detection system (chapter 5). We found that the optimised determinants yielded resolution values in the model that are indicated by a point on a trade-off curve. Setting the position of the trade-off point and the scale of the curve optimises the determinants as follows.

- **Light sensitivity** is optimised by scaling the curve as a function of mean illumination;
- **Contrast sensitivity** is optimised by scaling the curve as a function of average contrast;
- **Motion sensitivity** is optimised by moving the trade-off point along the curve as a function of the velocity of visual motion.
- **Heat production** is optimised by scaling the curve as a function of the capacity of the heat-decay system;
- **Energy consumption** is optimised by scaling the curve as a function of the capacity of the energy-supply system;

In summary, optimising the determinants yielded trade-off curves that determined the visual resolution of the model. Therefore, the determinants were confirmed to cause a trade-off between spatial and temporal resolution as is displayed in figure 7.1.

7.2 Additional determinants

In chapters 4 and 5 we found additional determinants that co-determine the optimal resolution values. For a full description of the additional determinants, we refer to sections 4.4, 4.5, 4.6, and 5.7. Here, we restrict ourselves to a summary. There are ten additional determinants: (1) eye size, (2) photoreceptor response threshold, (3) fraction between the width of the angular receptance function and the angular receptor spacing, (4) heat threshold, (5) heat production of individual receptors, (6) energy use of individual receptors, (7) noise level, (8) spatio-temporal pre-filter pass-band width, (9) closed-loop gain, and (10) detection threshold. For instance, tiny eyes are less sensitive to light than large eyes. Consequently, to obtain equal sensitivity, the optimal resolution of tiny eyes is lower than that of large eyes. Different animal species can have different values for the additional determinants. Hence, different species that live in similar ecological circumstances can have different optimal scaling factors of the trade-off curve. In conclusion, we expect that the additional factors to some extent influence the optimal scaling factor of the resolution of biological

Figure 7.1: The five determinants and a trade-off curve. Four determinants cause the curve to be scaled. One determinant causes that spatial resolution is traded for temporal resolution and *vice versa*.

visuo-motor systems. Future experiments should focus on the precise quantification of the influence.

7.3 Alternative occurrences of behaviour

Generic causes of a trade-off explain its existence in many occurrences of visually guided behaviour, *i.e.*, they are generalised causes (chapter 1). Below, we aim at generalising the determinants by discussing alternative occurrences of visually guided behaviour.

In chapter 3, we decided to model two occurrences (target tracking and gaze stabilisation) and their associated control systems (position-detection systems and velocity-detection systems, respectively). It was argued that the models address a large fraction, though not the complete set, of known occurrences of behaviour and control systems. For instance, in section 5.8.2, we mentioned that the response of a non-symmetric Reichardt correlator differs from that of our symmetric correlator. Therefore, non-symmetric Reichardt correlators may produce an alternative occurrence of gaze-stabilisation behaviour.

Three alternative occurrences are: target tracking and gaze stabilisation controlled by position-detection systems *and* velocity-detection systems; target tracking and gaze stabilisation controlled by systems based on vision and other types

of sensory input (*i.e.*, multimodally controlled behaviour); behaviours other than target tracking and gaze stabilisation. Each alternative is discussed in the respective subsections 7.3.1 to 7.3.3, followed by the generalisation of the determinants in subsection 7.3.4.

7.3.1 Combined position-detection and motion-detection systems

Many animal species employ the combined input of position and velocity information to track a moving target (Land, 1992). Moreover, the gaze-stabilisation and target-tracking systems interact to enable the occurrence of pursuit of moving objects against textured backgrounds (Kirschfeld, 1994). In chapters 4 to 5, we modelled the *isolated* behaviours of target tracking and gaze stabilisation using either position-detection or motion-detection systems. As a consequence, we expect that the determinants generalise to the aforementioned occurrence of behaviour as follows.

In chapters 4 and 5, the results show an inverse polynomial relation between the optimal values of spatial and temporal resolution. The general shape is described by $f_t^* = a \cdot (f_s^*)^b$ with a representing a scaling factor and b the polynomial coefficient. In the case of the target tracking model, b equals $-\frac{1}{2}$, whereas in the model of gaze stabilisation b equals -1 . We expect that the optimal resolution of combined position-detection and motion-detection systems lies on an interpolated version of the two curves. Alternatively, two separate pathways might exist that each have different resolution values (Livingstone and Hubel, 1988). Future research may clarify this issue. So far, we conclude that in general the determinants cause biological resolution to lie on trade-off curves that resemble those in chapters 4 and 5.

7.3.2 Multimodal control

A second alternative occurrence of behaviour, is the behaviour controlled by multiple sources of input, such as visual input and auditory input (Stein and Meredith, 1994). We provide two examples of such systems. First, the superior colliculus is a centre in the mammalian brain that integrates input from the eyes, ears, and tactile sensory organs to perform saccadic tracking (see, *e.g.*, Isa and Saito, 2001). Second, the semi-circular canals in the ears drive the vestibulo-ocular reflex when the head moves (Meiry, 1971) in coordination with the visually guided gaze-stabilisation system (Bosman, Ten Tusscher, and Kingma, 1999).

The two examples show that the behaviours studied in chapters 4 and 5 are not purely visually guided in most animals. Instead, vision is only important for the behaviour in the case of the smooth pursuit of slowly moving targets (Isa and Saito, 2001) and the stabilisation of gaze for slow oscillations of the head (*i.e.*, below 2 Hz, Carpenter, 1988). Therefore, the calculated trade-off curves best predict the optimal values of spatial and temporal resolution for slow visually guided behaviour. As a consequence, for instance, the optimal tuning velocity ($V^* = f_t^*/f_s^*$) for gaze

stabilisation may be lower than predicted, since non-visually controlled mechanisms control the behaviour at high input velocities.

We conclude that in many occurrence of behaviour, the determinant motion sensitivity is optimised to lower values than expected from our modelling results.

7.3.3 Control of other types of visually guided behaviour

The remaining alternative behaviours are those that do not involve target tracking nor gaze stabilisation. Below, we discuss three examples of such behaviours: colour vision, polarisation-based navigation, and active depth estimation.

Colour vision

Below we discuss how our determinants may be generalised to constitute the causes of the trade-off in colour vision. In our modelling studies, we focused on the brightness features of visual input (*i.e.*, light intensity and contrast). However, besides the detection of brightness, *i.e.*, the amplitude of electromagnetic waves, most animals detect colour, *i.e.*, the waves' frequency spectrum. In the following we examine briefly whether it is wise to study the trade-off in colour-sensitive and colour-blind parts of the brain separately, as we did. Thereafter, we determine whether the determinants can be generalised to colour vision.

In many insects, the regions of the eye and the brain dedicated to colour vision and those meant for target tracking and gaze stabilisation show little overlap (Stavenga, 2002). Therefore, in insects the resolution of colour-blind control systems can be studied separately from colour-vision systems. In the vertebrate retina, the colour-sensitive photoreceptors (cones) are mixed with the colour-blind receptors (rods; Wandell, 1995). This fact might pose problems to the justification of our experiments. However, upstream in the visual cortex, separate pathways exist for the detection of position, velocity, and colour (Livingstone and Hubel, 1988). The position-detection and velocity-detection pathways of both vertebrates and invertebrates are colour-blind (Menzel, 1979; Livingstone and Hubel, 1988). This fact suggests that the trade-off can be studied separately in position-detection, velocity-detection, and colour-detection systems.

Altogether, we tentatively conclude that the determinants presented in this thesis in particular explain a trade-off in those parts of the visual system that are colour-blind. Further work is necessary to understand the trade-off in colour vision.

Polarisation-based navigation

Another well-known type of visually guided behaviour is polarisation-based navigation. Due to scattering, sunlight is polarised in a regular pattern (Wehner, 1997). The orientation of the polarisation pattern depends on the position of the sun. Many animals use the pattern as a compass for navigation (Labhart, 1988; Dacke *et al.*, 2002). For instance, the dorsal rim of the eyes of desert ants and crickets is used

to detect the pattern of skylight polarisation (Möller *et al.*, 1998). The animals do not need high resolution for this task which generally results in lower values of spatial and temporal resolution than is expected from our results (Labhart, personal communication). As a consequence, we expect that our determinants apply to those parts of the eye and brain that are not devoted to polarisation-based navigation.

Active depth estimation

The third type of visually guided behaviour discussed here is the active estimation of depth, *e.g.*, through peering (Collett, 1978; Rind, 1997). When peering, an animal moves its head laterally to induce motion parallax on its retina (distant objects have slower retinal velocity than nearby objects). Using dedicated types of motion detectors, and the amplitude of the head movements, the distance to objects is estimated. We found no evidence in the literature that the head movements are guided by visual stimuli, *i.e.*, the behaviour is probably controlled in open-loop conditions. As a result, the gain-control properties of the Reichardt detector are not used in peering. In subsection 5.4.2 we determined that the gain-control properties of the Reichardt correlator set an upper limit to the tuning velocity and, therefore, to the visual resolution. From this result, we derive here that the resolution of the control system for active depth estimation is not matched to visual stimuli and is not likely to be constrained by a trade-off. Our determinants may, therefore, not be generalised to active depth estimation.

7.3.4 Consequences for the generalisation of the determinants

From subsections 7.3.1 to 7.3.3 we conclude that the experiments conducted in chapters 4 and 5 addressed the most important, though not all, occurrences of visually guided behaviour and their associated control systems. Based on this conclusion, we expect at least three consequences for the generalisation of our determinants to the visually guided behaviour of biological systems.

First, the determinants cause trade-off curves that are similar, though not exactly resemble the curves in our results. Second, the motion sensitivity (f_t^*/f_s^*) of visual systems that cooperate with other sensory-motor systems will probably appear lower than expected from the results. Third, parts of the brain that neither control brightness-based behaviour nor exploit the gain-control properties of the Reichardt correlators are likely not subject to a resolution trade-off caused by our determinants. Future research is required to verify the expectations.

7.4 Applying determinants to biological systems

To verify whether the five determinants meet the second criterion of generic causes of the trade-off, we aim at applying them to biological systems. For this, we compile a data set of biological resolution values in subsection 7.4.1. Subsection 7.4.2 predicts the trade-off points of all elements in the set. The actual trade-off points

are subsequently calculated and matched to the predictions in subsections 7.4.3 and 7.4.4. Finally, in subsection 7.4.5 we provide a summary.

7.4.1 A biological resolution data set

A single survey of spatial and temporal resolution values and of the data required to predict the trade-off point (optimal light sensitivity, optimal motion sensitivity, etc.) is not available (Warrant, personal communication). Therefore, we compiled our own set by combining several sources (Pick and Buchner, 1979; Mastebroek *et al.*, 1980; Holub and Morton-Gibson, 1981; Land, 1981; Anderson and Burr, 1985; O'Carroll *et al.*, 1996; O'Carroll *et al.*, 1997; Warrant, 1999). Our set contains the spatial and temporal resolution of 12 animal species and is displayed in table 7.1. Besides resolution values, we also include a brief description of each species' visual ecology (eye type, typical behaviour, and pattern of activity), from which we predict the trade-off point.

The resolution values are matched as follows. The location of a trade-off point (f_s^* , f_t^*) is determined by its position on a trade-off curve (f_t^*/f_s^* ; section 5.3) and the curve's scale ($f_s^* \cdot f_t^*$; section 5.5). Below we predict the ordering of the positions and scales of the trade-off points in the set. Thereafter, the actual positions and scales are calculated and matched to the predicted orderings.

7.4.2 Predicting trade-off points

The position on the trade-off curve sets motion sensitivity (section 5.3). Accordingly, we expect that animals that encounter high levels of image motion take high positions on the trade-off curves (*i.e.*, they are expected to have high values of the quotient f_t^*/f_s^*). The level of image motion depends on factors such as the stability of the behaviour and the velocity of prey targets. The ordering of positions is predicted to follow at least the two factors mentioned.

The scale of the trade-off curve determines the light sensitivity, contrast sensitivity, heat production, and energy consumption (section 5.5). According to Kleiber's law, the metabolic activity, *i.e.*, heat production and energy consumption, of large animals is higher than that of small animals (Kleiber, 1932). Therefore, we expect that body size is an important factor predicting the ordering of scales (*i.e.*, the value of the product $f_s^* \cdot f_t^*$). Moreover, the pattern of activity (day/night) sets the requirements for light sensitivity and contrast sensitivity. As a result, we predict that a diurnal animal's scale is ordered higher (*i.e.*, lower sensitivity) than that of a nocturnal animal. However, the light sensitivity is subject to many more determinants, such as eye size and eye type (subsection 7.2). Therefore, the additional determinants might rearrange the ordering of scale as predicted by the pattern of activity. Future research might specify a more precise ordering based on many more determinants.

Species	f_s^* (c/deg)	f_t^* (Hz)	Description
Bee-fly	0.1	6	compound eyes; hovers while feeding on nectar; diurnal
Beetle	0.2	20	superposition compound eyes; non-aerodynamic flyer; crepuscular (feeds in twilight);
Blowfly	0.1	10	compound eyes; large fly species; diurnal
Bumblebee	0.07	25	compound eyes; non-aerodynamic flyer, diurnal
Cat	0.5	5	single-lens eyes with reflecting layer; hunts prey from still position; nocturnal hunter
Housefly	0.05	2	neural superposition eyes; walks and flies steadily; diurnal
Hoverfly	0.2	5	compound eyes; hovers and performs fast chases; diurnal
Human	7	7	single-lens eye; optimised for slow tracking; diurnal
Locust	0.4	20	apposition compound eyes; walks and flies; diurnal
Moth (<i>Deilephila</i>)	0.05	0.7	compound eye; hovering; nocturnal
Moth (<i>Macroglossum</i>)	0.1	2	compound eye; hovering; diurnal
Toad	0.5	5	single-lens eyes; predates slowly moving animals; nocturnal

Table 7.1: Data set of spatial (f_s^*) and temporal (f_t^*) resolution values of different animal species. Also indicated are the brief descriptions of the visual ecology (eye type, typical behaviour, and pattern of activity). The ordering of species is alphabetical.

7.4.3 Matching position

The calculated positions (f_t^*/f_s^*) of the trade-off points are labelled in figure 7.2. Below we study whether the labels in the graph are ordered as is predicted by the

Figure 7.2: The position f_t^* / f_s^* of biological resolution values (table 7.1) on a trade-off curve. A high value of the fraction f_t^* / f_s^* indicates high motion sensitivity. A low values suggests sensitivity to low image motion.

two factors identified above (stability of the behaviour and velocity of prey targets). The top of the figure indicates that bumblebees and the selected beetle species (non-aerodynamic, *i.e.*, very unstable, flyers) are highly sensitive to motion. Below those, we find the species of which the behaviour is more stable than that of beetles and bumblebees: blowflies, beeflies, locusts, houseflies, and hoverflies fly more smoothly and sometimes hover in mid-air. Further towards the bottom of the figure, the two moth species are depicted which are true hoverers and, therefore, encounter even less

image motion. Cats and toads stay still while waiting for prey to appear, *i.e.*, they encounter little image motion during their hunting behaviour. Moreover, the target velocity of prey animals is rather slow (Warrant, 1999). Both facts might explain the low position of cats and toads in the figure. Human optomotor behaviour is optimised for lower image velocities than most insects (see subsection 4.7.1). Moreover, the human gaze-stabilisation system employs both visual and kinetic stimuli (subsection 7.3.4) which may explain why the human trade-off point has the lowest position in the figure. Altogether, the motion-sensitivity values presented above suggest that the calculated positions are matched to the ordering of positions predicted in subsection 7.4.2.

7.4.4 Matching scale

Figure 7.3 shows the labels of the calculated scales ($f_s^* \cdot f_t^*$) of the trade-off curves. In subsection 7.4.2, we predicted that the scales are ordered according to body size and according to the pattern of activity. First, the predicted ordering is, to a certain extent, reflected in the figure. For instance, the human trade-off point has a larger scale than that of all other, smaller species. Also the largest insect species (locust, beetle, bumblebee, and blowfly) are ordered according to body size. However, the scales of the remaining vertebrate species (cat and toad) are out of place among those of the much smaller insects. Moreover, the two moth species in the bottom of the figure do not follow the predicted ordering. The moths are larger than most of the flies. Hence, their scales are expected to exceed – not be exceeded by – those of the flies.

Second, the scales match the predicted ordering according to the pattern of activity when taking the additional determinants into consideration. In the figure we observe that the human and locust visual systems have the highest scales, *i.e.*, are least sensitive to light and contrast, which corresponds to their diurnal life style. In contrast, the nocturnal moth *Deilephila* has a low scale, *i.e.*, a high light sensitivity, which may support visually guided behaviour at night. Therefore, the extremes of the graph seem to match the predicted ordering well. However, we notice an irregularity with respect to our prediction in the centre of the graph. There, the nocturnal hunters, *i.e.*, toad and cat, as well as the crepuscular beetle, have a higher scale, *i.e.*, they seem to be less sensitive to light and contrast, than the diurnal insect species (bumblebee, blowfly, etc.). This irregularity may be put straight when taking into account three additional determinants (subsection 7.2). First, the eyes of toads and cats are of the single-lens type, which have a much larger aperture than compound eyes. Second, the signals of multiple photoreceptors (*viz.* rods) are pooled in single-lens eyes to produce a better signal. Third, the selected beetle species improves the visual signal optically. Consequently, the light sensitivities of single-lens eyes and superposition eyes are larger than predicted from the resolution values in our data set (cf. Warrant, 1999). From this fact we may derive that, when considering additional factors, the scales of the trade-off curves are matched to the predicted ordering by the pattern of activity.

Figure 7.3: The scale $f_s^* \cdot f_t^*$ trade-off curves that describe biological resolution values (see table 7.1). A high value of the product $f_s^* \cdot f_t^*$ indicates low sensitivity to light and contrast and high levels of heat production and energy consumption (low values of the product: *vice versa*).

7.4.5 Section summary

In this section we compiled a data set of biological resolution values in subsection 7.4.1. We calculated the positions and scales of the trade-off points of each animal in the set. In subsection 7.4.2 the orderings of positions and scales were predicted based on the causes of the trade-off revisited in section 7.1. We collected data of motion sensitivity which suggest that the trade-off points of the species in our data set match the predicted ordering of position (subsection 7.4.3). Moreover, although the calculated scales did not match the ordering by body size, it was hinted that the scale values match the ordering as predicted by the pattern of activity and

eye design (subsection 7.4.4). Based on our results in this section, we presume that the five factors that cause a trade-off in our situated models are determinants of the trade-off in biological systems as well.

7.5 Applying determinants to artificial systems

To further examine whether the five determinants are generic causes of the trade-off we suggest new ways in which our results can be applied to non-biological systems. In chapter 1 we already compared the trade-off between spatial and temporal resolution in biological visual systems to the trade-off found in photocameras. We found similar results in our models of visually guided behaviour. For instance, to obtain high light sensitivity the shutter time or the grain size of photosensitive films is increased (*i.e.*, the temporal or spatial resolution is decreased). Moreover, sharp images of fast moving objects in dark scenes are obtained by employing a short shutter time and a large grain size of the film, *i.e.*, the spatial resolution is traded for temporal resolution.

In addition to still devices, such as cameras, our result also apply to moving devices: situated artificial visuo-motor systems, such as eye prostheses, or exploration rovers. We showed that the visual resolution is of great importance to control the factors motion, light, and contrast sensitivity as well as the heat production and energy consumption of artificial visual devices. Hence, our results of the resolution trade-off are of particular importance to the design of miniature, low-power devices, for which all five factors are limiting.

Moreover, we showed that the resolution of visuo-motor systems based on Reichardt correlators determines the built-in gain control of those systems. Exploiting the properties of Reichardt correlators in artificial devices will provide a cheap and robust solution to stability problems. Finally, we determined that the resolution scale sets both the SNR and latency of visuo-motor systems to optimise gaze-stabilisation behaviour. Therefore, our results help determine the optimal resolution of situated artificial visuo-motor systems that perceive and act in real-time circumstances.

Altogether, we showed above that the determinants of the trade-off in our situated models may cause a trade-off in non-biological applications as well.

7.6 The generic causes

In this chapter, we iterated the five determinants (section 7.1). Then our choices from chapter 3, concerning determinants and model behaviours, were evaluated in sections 7.2 and 7.3. Finally, sections 7.4 and 7.5 applied the determinants to biological systems and to artificial systems, respectively.

Based on our results we determine that the five determinants selected in chapter 3 comply to the three conditions in our definition of generic causes in chapter 1. (1) They explain the existence of a trade-off in many occurrences of visually guided

behaviour; (2) they cause a trade-off in many different biological visual systems; and (3) they cause a trade-off in artificial visual systems.

Therefore, we may conclude that the five determinants light sensitivity, contrast sensitivity, motion sensitivity, heat production, and energy consumption are generic causes of the trade-off between spatial and temporal resolution in visually guided behaviour. By comparing biological resolution to the predicted trade-off points in our models, we remark that there exist additional determinants, such as eye design (subsection 7.4.4). Future research is required to determine the influence of the additional determinants on the trade-off.

Chapter 8

Conclusions

*“There yet remains but one concluding tale,
And then this chronicle of mine is ended.”*

Alexander Pushkin (1799 – 1837), Russian poet.

Chapter contents: Chapter conclusions summarised, generic causes of the trade-off, situated modelling: a general conclusion, future research

Our research is rounded up by drawing conclusions from it. We start by summarising the conclusions from various earlier chapters (section 8.1). Then we draw two main conclusions from our research in sections 8.2 and 8.3. Finally, section 8.4 describes future research directions.

8.1 Chapter conclusions summarised

In chapters 2 to 7 a series of chapter conclusions was drawn. Below we summarise the highlights.

Chapter 2 All approaches to studying visually guided behaviour were categorised in a taxonomy which comprised three dimensions: the level of detail, the class of research, and the type of investigation. We expected that an approach of the systems level, of the modelling class, and of the situated type would be the most appropriate approach to answering our research question. Therefore, it was concluded to employ situated systems modelling for our research.

Chapter 3 Based on our analysis and selection of occurrences of visually guided behaviour and of the determinants of the trade-off, we developed the following method to scrutinise the trade-off. In the method the determination of

the causes of the trade-off was accomplished by studying five selected determinants of the trade-off in two situated models of visually guided behaviour (*i.e.*, the five determinants light sensitivity, contrast sensitivity, motion sensitivity, heat production, and energy consumption and two models based on either a position-detection system or a motion-detection system). In conclusion, this method was expected to provide satisfactory answers to our research question.

Chapter 4 Our modelling study of position-detection systems yielded two conclusions. First, four of the five determinants definitely caused a trade-off between spatial and temporal resolution in our model of target-tracking behaviour (light sensitivity, motion sensitivity, heat production, and energy consumption). A fifth factor (contrast sensitivity) was not investigated, but was expected to yield the same results. Second, by analogy we tentatively concluded that the four determinants caused a trade-off in the resolution values of biological position-detection systems as well.

Chapter 5 From the results of this chapter we were able to draw a series of eight conclusions (subsection 5.8.3). We repeat the two most important ones here. First, the factors motion sensitivity, light sensitivity, and contrast sensitivity caused a trade-off between the spatial and temporal resolution of our model of motion-detection systems. The factors heat production and energy consumption were not investigated but were expected as well to cause a trade-off in motion-detection systems. Second, we concluded that the determinants caused a trade-off in both artificial and biological visual systems. In summary, the conclusions of chapter 5 generalised those of chapter 4.

Chapter 6 A qualitative comparison of our results of situated systems modelling to those of controlled behavioural studies revealed the advantages of our approach. Similarly, a quantitative comparison showed that our results of situated systems modelling were more plausible than those of information theory. Based on the two comparisons, we stated that our approach outperformed its two main alternative approaches. In conclusion, we stated that situated systems modelling was vital to studying the trade-off.

Chapter 7 We verified that our experiments in chapters 4 and 5 covered the main occurrences of visually guided behaviour. Moreover, the determinants explained the existence of a trade-off in biological, as well as artificial visual systems. Therefore, we concluded that the five determinants selected in chapter 3 were generic causes of the trade-off between spatial and temporal resolution.

8.2 Generic causes of the trade-off

To draw conclusions on the generic causes of the trade-off we start by quoting Land (1981) who observed:

“It is commonly pointed out that insects have poor spatial resolution because of the optical limitations of compound eyes, but it is almost certainly true that in the rapid flyers [...] the equation is tipped in favour of temporal resolution.” (p. 480)

In our research we have explained this observation by showing that the spatial and temporal resolution of visual systems were subject to a trade-off. The most appropriate research approach to studying the trade-off addressed the optimisation of visual resolution in the context of the visual ecology of animals. Situated models of visually guided behaviour served as good platforms for such optimisation studies. They led us to new insights into the constraints and optimality of biological visuo-motor systems.

In chapter 1 we formulated the main question of our research as: “which are the generic causes of the trade-off between spatial and temporal resolution in visually guided behaviour?” From our results we derived the following conclusion. The five determinants (light sensitivity, contrast sensitivity, motion sensitivity, heat production, and energy consumption) constituted the main generic causes of the trade-off. This conclusion is based on three assumptions that were verified in chapter 7. First, the determinants explained the existence of a trade-off in many occurrences of visually guided behaviour, such as those controlled by position-detection systems, by motion-detection systems, and by combined systems that detect the position *and* motion in a visual scene. Second, the determinants predicted the trade-off points of animals as a function of their visual ecology. Third, the determinants provided engineering principles for the construction of low-power, artificial, visuo-motor systems, such as eye prostheses. Therefore, our results provided an important contribution to the study of biological and artificial visual systems.

8.3 Situated modelling: a general conclusion

The strength of our approach of situated modelling appeared from qualitative and quantitative comparisons to its two main alternative approaches (controlled behavioural studies and information theory, respectively). First, situated modelling solved some of the problems of controlled behavioural studies, such as those with many different measures, those with a limited availability of data, and those with a lack of insight. Second, our approach yielded quantitative results that were more plausible from an evolutionary point of view than the results of information theory with a similar model. Therefore, the use of situated models outperformed the alternative approaches in finding the generic causes of the trade-off. As a general

conclusion on the research approach, we state that situated modelling was vital to studying the trade-off between spatial and temporal resolution.

8.4 Future research

In the previous chapters, we have hinted at many lines for future research. Below we summarise the three most important ones.

First, we suggest to employ mobile robots to study the trade-off (cf. Krichmar and Edelman, 2002). The use of robots will increase the ecological validity of situated models owing to the emulation of naturalistic behaviour of animals in their actual natural environment (Webb, 2002). Second, in section 7.2 we summarised ten additional determinants of the trade-off. Further investigation of their influence on the optimal resolution values of both situated models and of animals will increase our understanding of the trade-off. Third, section 7.3 summarised alternative occurrences of visually guided behaviour that were not investigated in our studies. For instance, the possible existence of a trade-off in colour vision has not been verified until now. Future research on alternative occurrences of behaviour is expected to lead to a better generalisation of our results.

Altogether, we state that our research and its scientific implications have accumulated into better understanding of animals and the ways that they act in a visual world. Although there remain many questions to investigate, our conclusions have taken us a step further towards the resolution of visually guided behaviour.

Appendix A

Filter response functions

In chapter 5 we employed three types of analogue filters: Gaussian spatial low-pass filters; first-order temporal low-pass filters; and first-order temporal high-pass filters. In tables A.1 to A.3 we characterise each filter type by providing their response functions.

Analogue filters are generally characterised by four response functions (Oppenheim *et al.*, 1997). First, the impulse response h denotes the filter response to the Dirac delta-function δ in the spatial or temporal domain. The impulse response of a spatial filter is denoted by $h(x)$; that of a temporal filter by $h(t)$. Second, $H(i\omega)$ is the Fourier transformed impulse response function and describes the filter response in the frequency domain as a complex value. The variable ω is the angular frequency of an input signal defined as $\omega = 2\pi \cdot f$ (with f representing either the spatial frequency f_s or the temporal frequency f_t). Third, the amplitude response $|H(\omega)|$ is the absolute value of the frequency response and determines the modulation transfer of the filter. Finally the phase response $\angle H(\omega)$ is the phase angle of the frequency response and denotes the temporal or spatial phase shift as a function of the angular frequency.

In the tables A.1 to A.3 we used the variables as defined above and the following additional variables: $\Delta\sigma$ is the Gaussian's half-width; τ is a time-constant of the temporal filters; and $\epsilon(t)$ is the unity step function.

	Gaussian spatial low-pass filter
Impulse response	$h(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2(\Delta\sigma)^2}\right)$
Frequency response	$H(i\omega) = \exp(-\frac{1}{2}(\Delta\sigma)^2\omega^2)$
Amplitude response	$ H(\omega) = H(i\omega)$
Phase response	$\angle H(\omega) = 0$

Table A.1: The response of a Gaussian spatial low-pass filter.

	First-order temporal low-pass filter
Impulse response	$h(t) = \frac{1}{\tau} \cdot \exp(-t/\tau) \cdot \varepsilon(t)$
Frequency response	$H(i\omega) = \frac{1}{i\omega\tau+1}$
Amplitude response	$ H(\omega) = \frac{1}{\sqrt{1+(\omega\tau)^2}}$
Phase response	$\angle H(\omega) = -\arctan(\omega\tau)$

Table A.2: The response of a first-order temporal low-pass filter.

	First-order temporal high-pass filter
Impulse response	$h(t) = \frac{1}{\tau} \cdot \exp(-t/\tau) + \delta(t) \cdot \varepsilon(t)$
Frequency response	$H(i\omega) = \frac{i\omega\tau}{i\omega\tau+1}$
Amplitude response	$ H(\omega) = \frac{\omega\tau}{\sqrt{1+(\omega\tau)^2}}$
Phase response	$\angle H(\omega) = \frac{1}{2}\pi - \arctan(\omega\tau)$

Table A.3: The response of a first-order temporal high-pass filter.

Appendix B

Stability of the update rule

In this appendix, we set the conditions for stability of the update procedure used in chapter 5. For this, we calculate the bounding conditions of possible values of the learning parameter η that guarantee convergence of the error value to zero. If convergence to zero is theoretically impossible, we propose a strategy to find the minimal error.

To calculate the bounding conditions, we consider an arbitrary value of K and determine the corresponding open-loop gain g and error value e . After updating K , the new open-loop gain g' is expressed as $g' = (K_c + \Delta K_c) \cdot (K + \Delta K)$ (in steady-state conditions, the mean integrator gain is constant, *i.e.*, $K_i = 1$ and $\Delta K_i = 0$). The updated error equals $e' = g^* - g'$. To check for convergence, we first rewrite e' in terms of e :

$$\begin{aligned} e' = g^* - g' &= g^* - (K_c + \Delta K_c) \cdot (K + \Delta K) \\ &= g^* - K_c \cdot (K + \eta \cdot \frac{e}{g} \cdot K) - \Delta K_c \cdot (K + \Delta K) \\ &= g^* - g - \eta \cdot e - \Delta K_c \cdot (K + \Delta K) \\ &= e \cdot (1 - \eta) - \Delta K_c \cdot (K + \Delta K) \end{aligned} \tag{B.1}$$

Second, using equation (B.1), we express the error as a function of discrete time t :

$$e(t) = (1 - \eta) \cdot e(t-1) - \Delta K_c \cdot (K + \Delta K) \tag{B.2}$$

If we assume that the value of K_c does not change with K , *i.e.*, $\Delta K_c = 0$, the limit of equation (B.2) yields $\lim_{t \rightarrow \infty} e(t) = 0$ iff $|1 - \eta| < 1$. Therefore, when the assumption holds, the bounding conditions are $0 < \eta < 2$.

However, we shown that ΔK_c depends on K . To indicate the true bounding conditions of η , we calculate g and K_c for the values of K and $\overline{V}_r$ shown in figure 5.15 (a). The results are shown in figure B.1, maximum normalised. In the figure, we distinguish three regions, in each of which we inspect the derivatives of

Figure B.1: The open-loop gain g (solid curve) and correlator gain K_c (dashed curve), both normalised, as a function of K in closed-loop experiments. The significance of region 1 to 3 is described in the main text.

K_c and g with respect to K . First, for $K \in [0, K_1]$, $\frac{dK_c}{dK} > 0$ and $\frac{dg}{dK} > 0$; second, for $K \in [K_1, K_2]$, $\frac{dK_c}{dK} < 0$ and $\frac{dg}{dK} > 0$; third, for $K \in [K_2, \infty)$, $\frac{dK_c}{dK} < 0$ and $\frac{dg}{dK} < 0$.

Equation (5.22) shows that the sign of ΔK equals that of e . From this fact and the derivatives mentioned above, we derive that in the first region, $\text{sign}(\Delta K_c) = \text{sign}(e)$, whereas in the second and third region, the signs are opposite. Therefore, in the first region, equation (B.2) is approximated as $e(t) = (1 - \eta') \cdot e(t-1)$, with $\eta' > \eta$, i.e., in the first region the effective learning rate η' is larger than the parameter η , since ΔK_c boosts the change in g . As a consequence, the bounding conditions of η are decreased to guarantee convergence. In contrast, in the second region, the effective learning rate η' has a smaller value than the parameter η , since ΔK_c and ΔK have opposite signs (ΔK_c reduces the change in g). Therefore, in the second region, the bounding conditions of η are increased to guarantee convergence. Altogether, in the first and second region, both the upper and lower bounding conditions of η are drawn towards the value 1.

In the third region, not only ΔK_c and ΔK have opposite signs, also the gain g decreases when K is increased. Since e and ΔK have the same sign, the error e is, therefore, increased when updating K according to equation (5.22) in the third region (ΔK_c overrides the influence of ΔK). Therefore, the learning parameter η should be negative to let the error converge in the third region.

Concluding, there exist conflicting bounding conditions for the three regions. Moreover, the bounds of the regions themselves are not constant, but vary with

the tuning velocity of the correlator. Therefore, we apply the following solution to optimise the value of K for any tuning velocity and background velocity. We set the learning parameter $\eta \approx 1$. If the gain factor K is initialised to a value in the first or second region, the error e converges to 0 as explained before. If, however, the target gain g^* exceeds the maximal gain $g_{\max}$ for the current combination of tuning velocity and background velocity, K is updated into the third region, which leads to divergence of the error value for $\eta \approx 1$. To make the process finite, we store the minimal error value and the corresponding value of K during updating and stop iterating when K surpasses a threshold θ_K . When this threshold is met, we return the K value that produces the lowest error value found.

References

- Adelson, E.H. and Bergen, J. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the optical society of America A*, Vol. 2, pp. 284–299. [25]
- Alcock, J. (1993). *Animal behavior*. Sinauer Associates, Sunderland, MA, fifth edition. [24]
- Anderson, S.J. and Burr, D.C. (1985). Spatial and temporal selectivity of the human motion detection system. *Vision research*, Vol. 25, No. 8, pp. 1147–1154. [12, 116, 133]
- Arbib, M.A. (1987). Levels of modeling of mechanisms of visually guided behaviour. *Behavioural brain science*, Vol. 10, pp. 407–465. [10, 17, 24]
- Arbib, M. (ed.) (1995). *The handbook of brain theory and neural networks*, pp. 46–48. MIT Press, Cambridge, MA. [15]
- Arnett, D.W. (1972). Spatial and temporal integration properties of units in first optic ganglion of Dipterans. *Journal of neurophysiology*, Vol. 35, pp. 429–444. [25, 67]
- Bäck, T. (1996). *Evolutionary algorithms in theory and practice*. Oxford University Press, New York, NY. [45]
- Baddeley, R. (2000). Introductory information theory and the brain. *Information theory and the brain* (eds. R. Baddeley, P. Hancock, and P. Földiák), pp. 1–19, Cambridge University Press, Cambridge. [18, 19]
- Baddeley, R., Hancock, P., and Földiák, P. (eds.) (2000). *Information theory and the brain*. Cambridge university press, Cambridge. [18]
- Baker, M.A. (1979). A brain-cooling system in mammals. *Scientific American*, Vol. 240, pp. 130–139. [32, 53, 59]
- Baker, M.A. and Chapman, L.W. (1977). Rapid brain cooling in exercising dogs. *Science*, Vol. 195, No. 4280, pp. 781–783. [32]

- Ballard, D.H. (1991). Animate vision. *Artificial intelligence*, Vol. 48, pp. 57–86. [20]
- Barlow, H.B. (1952). The size of ommatidia in apposition eyes. *Journal of experimental biology*, Vol. 29, pp. 667–674. [12]
- Barlow, H.B. (1956). Retinal noise and absolute threshold. *Journal of the optical society of America*, Vol. 46, pp. 634–639. [30, 83]
- Barlow, H.B. (1961). Possible principles underlying the transformations of sensory messages. *Sensory communication* (ed. W.A. Rosenblith), pp. 217–234, MIT Press, Cambridge, MA. [15, 19]
- Barlow, H.B. (1981). Critical limiting factors in the design of the eye and visual cortex. *Proceedings of the Royal Society of London B*, Vol. 212, pp. 1–34. [6, 27]
- Barlow, H.B. and Levick, W.R. (1965). The mechanism of directionally selective units in rabbit's retina. *Journal of physiology*, Vol. 178, pp. 477–504. [25]
- Beersma, D.G.M., Stavenga, D.G., and Kuiper, J.W. (1977). Retinal lattice, visual field, and binocularities in flies. *Journal of comparative physiology*, Vol. 119, pp. 207–220. [33]
- Blanchard, M., Rind, F.C., and Verschure, P.F.M.J. (2001). How accurate need sensory coding be for behaviour? Experiments using a mobile robot. *Neurocomputing*, Vol. 38-40, pp. 1113–1119. [2]
- Borst, A. and Egelhaaf, M. (1989). Principles of visual motion detection. *Trends in neuroscience*, Vol. 12, No. 8, pp. 297–306. [25, 26]
- Borst, A. and Egelhaaf, M. (1990). Direction selectivity of blowfly motion-sensitive neurons is computed in a two-stage process. *Proceedings of the national academy of science of the USA*, Vol. 87, pp. 9363–9367. [26]
- Bosman, J., Ten Tusscher, M.P.M., and Kingma, H. (1999). The influence of visual input on the vestibulo-ocular reflex. *Documenta ophthalmologica*, Vol. 99, pp. 83–92. [130]
- Buchner, E. (1976). Elementary movement detectors in an insect visual system. *Biological cybernetics*, Vol. 24, pp. 85–101. [71]
- Burton, B.G. (2000). Problems and solutions in early visual processing. *Information theory and the brain* (eds. R. Baddeley, P. Hancock, and P. Földiák), pp. 25–40, Cambridge University Press, Cambridge. [18]
- Cabanac, M. and Brinnet, H. (1985). Blood flow in the emissary veins of the human head during hypothermia. *European journal of applied physiology*, Vol. 54, pp. 172–176. [53, 59]

- Carpenter, R.H.S. (1988). *Movements of the eye*. Psion Limited, London. [130]
- Clancey, W.J. (1997). *Situated cognition: on human knowledge and computer representations*. Cambridge University Press, New York, NY. [20]
- Clark, A. (1996). *Being there: putting brain, body, and world together again*. MIT Press, Cambridge, MA. [19, 20]
- Cliff, D. (1995). Computational neuroethology. *The handbook of brain theory and neural networks*, pp. 626–630, MIT press, Cambridge, MA. [24]
- Collett, T.S. (1978). Peering — a locust behaviour pattern for obtaining motion parallax information. *Journal of Experimental Biology*, Vol. 76, pp. 237–241. [132]
- Collett, T.S. (1980). Angular tracking and the optomotor response – an analysis of visual reflex interaction in a hoverfly. *Journal of comparative physiology A*, Vol. 140, pp. 145–158. [2, 11, 24, 25, 39, 83, 91]
- Collett, T.S. and Land, M.F. (1975). Visual control of flight behaviour in the hoverfly, *Syrphia pipiens* L. *Journal of comparative physiology A*, Vol. 99, pp. 1–66. [6, 37, 39, 44, 58, 60]
- Crawford, M.L.J., Andersen, R.A., Blake, R., Jacobs, G.H., and Neumeyer, Ch.N. (1990). Interspecies comparison in the understanding of human visual perception. *Visual perception — the neurophysiological foundations* (eds. L. Spillmann and J.S. Werner), pp. 23–52, Academic Press, San Diego, CA. [116]
- Dacke, M., Nordström, P., Scholtz, C.H., and Warrant, E.J. (2002). A specialized dorsal rim area for polarized light detection in the compound eye of the scarab beetle *Pachysoma striatum*. *Journal of comparative physiology*, Vol. 188, pp. 211–216. [131]
- De Lange, F., Bovee-Geurts, P.H.M., Van Oostrum, J., Portier, M.D., Verdegem, P.J.E., Lugtenberg, J., and De Grip, W.J. (1998). An additional methyl group at the 10-position of retinal dramatically slows down the kinetics of the rhodopsin photocascade. *Biochemistry*, Vol. 37, pp. 1411–1420. [16]
- Dror, R.O., O’Carroll, D.C., and Laughlin, S.B. (2001). Accuracy of velocity estimation by Reichardt correlators. *Journal of the optical society of America A*, Vol. 18, No. 2, pp. 241–252. [64, 66, 67, 72, 73, 116, 117]
- Edwards, C.H. and Penney, D.E. (1982). *Calculus and Analytic Geometry*, Chapter Chapter 7. Exponential and Logarithmic Functions, pp. 344–389. Prentice-Hall, Englewood Cliffs, NJ. [53]
- Egelhaaf, M. and Borst, A. (1989). Transient and steady-state response properties of movement detectors. *Journal of the optical society of America*, Vol. 6, pp. 116–127. [68]

- Egelhaaf, M. and Borst, A. (1993a). A look into the cockpit of the fly: visual orientation, algorithms, and identified neurons. *The journal of neuroscience*, Vol. 13, No. 11, pp. 4563–4574. [9, 10, 11, 17, 61]
- Egelhaaf, M. and Borst, A. (1993b). Movement detection in arthropods. *Visual motion and its role in the stabilization of gaze* (eds. F.A. Miles and J. Wallman), Vol. 5 of *Reviews of oculomotor research*, pp. 53–77, Elsevier, Amsterdam. [65]
- Egelhaaf, M., Borst, A., and Reichardt, W. (1989). Computational structure of a biological motion-detection system as revealed by local detector analysis in the fly's nervous system. *Journal of the optical society of America*, Vol. 6, pp. 1070–1087. [68, 69]
- Egelhaaf, M., Grewe, J., Kern, R., and Warzecha, A-K. (2001). Outdoor performance of a motion-sensitive neuron in the blowfly. *Vision research*, Vol. 41, pp. 3627–3637. [15]
- Egelhaaf, M., Kern, R., Krapp, H.G., Kretzberg, J., Kurtz, R., and Warzecha, A-K. (2002). Neural encoding of behaviourally relevant visual-motion information in the fly. *Trends in Neurosciences*, Vol. 25, No. 2, pp. 96–102. [15]
- Eizenman, M., Hallett, P.E., and Frecker, R.C. (1985). Power spectra for ocular drift and tremor. *Vision research*, Vol. 25, No. 11, pp. 1635–1640. [83, 87]
- Euler, Th., Detwiler, P.B., and Denk, W. (2002). Directionally selective calcium signals in dendrites of starburst amacrine cells. *Nature*, Vol. 418, pp. 845–852. [63]
- Falk, D. (1992). Evolution of the brain and cognition in hominids. *The sixty-second James Arthur lecture*, The American museum of natural history, New York, NY. [32]
- Falk, D. and Gage, T.B. (1998). Radiators are cool: a response to Braga and Boesch's published paper and reply. *Journal of human evolution*, Vol. 35, pp. 307–312. [32, 53]
- Felleman, D.J. and Van Essen, D.C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral cortex*, Vol. 1, pp. 1–47. [14]
- Feng, C. and Plamondon, R. (2001). On the stability analysis of delayed neural networks systems. *Neural networks*, Vol. 14, pp. 1181–1188. [86]
- Fermi, G. and Reichardt, W. (1963). Optomotorische Reaktionen der Fliege *Musca domestica*. *Kybernetik*, Vol. 2, pp. 15–28. [12]
- Fernald, R.D. (2000). Evolution of eyes. *Current opinion in neurobiology*, Vol. 10, pp. 444–450. [1, 16, 27]

- Field, D.J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the optical society of America A*, Vol. 4, pp. 2379–2394. [73]
- Fiorentini, A., Baumgartner, G., Magnussen, S., Schiller, P.H., and Thomas, J.P. (1990). The perception of brightness and darkness. *Visual perception – the neurophysiological foundations* (eds. L. Spillmann and J.S. Werner), pp. 129–161, Academic press, San Diego, CA. [43, 67]
- Floreano, D. and Mondada, F. (1996). Evolution of homing navigation in a real mobile robot. *IEEE transactions on systems, man, and cybernetics*, Vol. 26, No. 3, pp. 396–407. Part B: cybernetics. [41]
- Franceschini, N., Pichon, J-M., and Blanes, C. (1992). From insect vision to robot vision. *Philosophical transactions of the Royal Society of London B*, Vol. 337, No. 1281, pp. 283–294. [20]
- Fritzsche, B. (1995). Evolution of ancestral vertebrate brains. *The handbook of brain theory and neural networks* (ed. M.A. Arbib), pp. 373–377, MIT Press, Cambridge, MA. [15]
- Geissler, W.S. and Albrecht, D.G. (2000). Spatial vision. *Seeing - handbook of perception and cognition (2nd edition)* (ed. K.K. De Valois), pp. 79–128, Academic Press, San Diego, CA. [61]
- Gibson, J.J. (1979). *The ecological approach to perception*. Houghton Mifflin, Boston, MA. [11, 14]
- Gilbert, C. and Strausfeld, N.J. (1991). The functional organization of male-specific visual neurons in flies. *Journal of comparative physiology A*, Vol. 169, pp. 395–411. [39]
- Goldberg, D.E. (1989). *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley, Reading, MA. [45]
- Götz, K.G. (1975). The optomotor equilibrium of the *Drosophila* navigation system. *Journal of comparative physiology*, Vol. 99, pp. 187–210. [12, 14]
- Hamdorf, K. (1979). The physiology of invertebrate visual pigments. *Handbook of sensory physiology* (ed. H. Autrum), Vol. VII/6A, pp. 145–224, Springer-Verlag, Berlin. [16]
- Hardie, R.C. (1985). Functional organization of the fly lamina. *Progress in sensory physiology* (ed. D. Ottoson), Vol. 5, pp. 1–80, Springer-Verlag, Berlin. [115]
- Harrison, R.R. and Koch, Ch. (2000). A robust analog VLSI Reichardt motion sensor. *Analog integrated circuits and signal processing*, Vol. 24, pp. 213–229. [25]

- Hassenstein, B. and Reichardt, W. (1956). Systemtheoretische Analyse der Zeit-, Reihenfolgen- und Vorzeichenauswertung bei der Bewegungsperception der Rüsselkäfers, *Chlorophanus*. *Zeitschrift für Naturforschung*, Vol. 11b, pp. 513–524. [13, 25]
- Hausen, K. (1982). Motion sensitive interneurons in the optomotor system of the fly. *Biological cybernetics*, Vol. 45, pp. 143–156. [14, 117]
- Heisenberg, M. and Wolf, R. (1984). *Vision in Drosophila*. Springer-Verlag, Berlin. [65, 66]
- Heisenberg, M. and Wolf, R. (1993). The sensory-motor link in motion-dependent flight control of flies. *Visual Motion and its Role in the Stabilization of gaze* (eds. J. Wallman and F.A. Miles), pp. 265–282, Elsevier, Amsterdam. [14]
- Hennig, M.H. and Funke, K. (2001). A biophysically realistic simulation of the vertebrate retina. *Neurocomputing*, Vol. 38-40, pp. 659–665. [15, 20]
- Hertz, J., Krogh, A., and Palmer, R.G. (1991). *Introduction to the theory of neural computation*. Addison-Wesley, Reading, MA. [40, 89]
- Holland, J.H. (1975). *Adaptation in natural and artificial systems*. University of Michigan Press, Ann Arbor, MI. [45]
- Holub, R.A. and Morton-Gibson, M. (1981). Response of visual cortical neurons of the cat to moving sinusoidal gratings: response-contrast functions and spatiotemporal interactions. *Journal of neurophysiology*, Vol. 46, No. 6, pp. 1244–1259. [116, 133]
- Huber, S.A. and Bülthoff, H.H. (1998). Simulation and robot implementation of visual orientation behaviors of flies. *From animals to animats 5, proceedings of the fifth international conference on simulation of adaptive behavior* (eds. R. Pfeifer, B. Blumberg, J.-A. Meyer, and S.W. Wilson), pp. 77–85, MIT Press, Cambridge, MA. [65]
- Isa, T. and Saito, Y. (2001). The direct visuo-motor pathway in the mammalian superior colliculus; novel perspective on the laminar connection. *Neuroscience research*, Vol. 41, pp. 107–113. [130]
- Jansonius, N.M. (1990). Properties of the sodium pump in the blowfly photoreceptor cell. *Journal of comparative physiology A*, Vol. 167, pp. 461–467. [16]
- Jerison, H.J. (1985). Animal intelligence as encephalization. *Philosophical transactions of the Royal Society of London B*, Vol. 308, pp. 21–35. [32]
- Juusola, M. and Hardie, R.C. (2001). Light adaptation in *Drosophila* photoreceptors: II. Rising temperature increases the bandwidth of reliable signaling. *Journal of general physiology*, Vol. 117, pp. 27–41. [115]

- Kalaska, J.F. (1995). Reaching movements: implications of connectionist models. *The handbook of brain theory and neural networks*, pp. 788–793, MIT press, Cambridge, MA. [24]
- Kaneko, A. and Shimazaki, H. (1976). Synaptic transmission from photoreceptors to the second-order neurons in the carp retina. *Neural principles in vision* (eds. F. Zettler and R. Weiler), pp. 143–157, Springer-Verlag, Berlin. [16]
- Keller, K. (ed.) (1993). *Science and technology of photography*. VCH, Weinheim. [6]
- Kern, R., Lutterklas, M., Petereit, Ch., Lindemann, J.P., and Egelhaaf, M. (2001a). Neuronal processing of behaviourally generated optic flow: experiments and model simulations. *Network: computation in neural systems journal*, Vol. 12, pp. 351–369. [12, 116]
- Kern, R., Petereit, Ch., and Egelhaaf, M. (2001b). Neural processing of naturalistic optic flow. *The journal of neuroscience*, Vol. 21:RCI39, pp. 1–5. [15, 73]
- Kety, S.S. (1957). The general metabolism of the brain in vivo. *Metabolism of the nervous system* (ed. D. Richter), pp. 221–237, Pergamon Press, London. [34, 57]
- Kirschfeld, K. (1969). Optics of the compound eye. *Processing of optical data by organisms and by machines* (ed. W. Reichardt), pp. 144–166, Academic Press, New York, NY. [42]
- Kirschfeld, K. (1976). The resolution of lens and compound eyes. *Neural principles in vision* (eds. F. Zettler and R. Weiler), Springer-Verlag, Berlin. [2, 4, 5, 42]
- Kirschfeld, K. (1994). Tracking of small objects in front of a textured background by insects and vertebrates: phenomena and neural basis. *Biological cybernetics*, Vol. 70, pp. 407–415. [24, 37, 38, 39, 130]
- Kleiber, M. (1932). Body size and metabolism. *Hilgardia*, Vol. 6, pp. 315–353. [133]
- Kortmann, R. (2001). Embodied cognitive science. *Proceedings of Robo Sapiens the first Dutch symposium on embodied intelligence* (eds. W. De Back, T. Van der Zant, and L. Zwanepol), Vol. 24 of *Artificial intelligence preprint series*, Universiteit Utrecht, Utrecht. [19]
- Kortmann, R. and Hallam, J.C.T. (1999). Studying animals through artificial evolution: the cricket case. *Lecture notes in artificial intelligence – Advances in artificial life* (eds. D. Floreano, J-D. Nicoud, and F. Mondada), Vol. 1674, pp. 215–224, Springer-Verlag, Berlin. [41]

- Kortmann, R., Postma, E.O., and Van den Herik, H.J. (1999). The trade-off between spatial and temporal resolution in visual systems. *Proceedings of the eleventh Belgian-Dutch conference on artificial intelligence* (eds. E.O. Postma and M. Gyssens), pp. 219–226. [27, 32]
- Kortmann, R., Postma, E.O., and Van den Herik, H.J. (2000). Visual resolution evolving to a trade-off curve. *From animals to animats 6, proceedings of the sixth international conference on simulation of adaptive behavior* (eds. J.-A. Meyer, A. Berthoz, D. Floreano, H. Roitblat, and S.W. Wilson), pp. 405–412, MIT Press, Cambridge, MA. [31]
- Kortmann, R., Postma, E.O., and Van den Herik, H.J. (2001). Evolution of visual resolution constrained by a trade-off. *Artificial life*, Vol. 7, No. 2, pp. 125–145. [37]
- Krauzlis, R.J. and Stone, L.S. (1999). Tracking with the mind's eye. *Trends in neuroscience*, Vol. 22, No. 12, pp. 544–550. [38]
- Krichmar, J.L. and Edelman, G.M. (2002). Machine psychology: autonomous behavior, perceptual categorization and conditioning in a brain-based device. *Cerebral cortex*, Vol. 12, pp. 818–830. [20, 144]
- Kuffler, S.W., Nichols, J.G., and Martin, A.R. (1984). *From neuron to brain: a cellular approach of the function of the nervous system*. Sinauer Associates, Sunderland, MA, 2nd edition. [16]
- Labhart, Th. (1988). Polarization-opponent interneurons in the insect visual system. *Nature*, Vol. 331, pp. 435–437. [131]
- Land, M.F. (1981). Optics and vision in invertebrates. *Handbook of sensory physiology* (ed. H. Autrum), Vol. VII/6B, pp. 471–592, Springer-Verlag, Berlin. [2, 3, 4, 27, 40, 66, 115, 133, 143]
- Land, M.F. (1992). Visual tracking and pursuit: humans and arthropods compared. *Journal of insect physiology*, Vol. 38, No. 12, pp. 939–951. [24, 25, 37, 39, 65, 66, 72, 74, 85, 96, 130]
- Land, M.F. (1997). Visual acuity in insects. *Annual review of entomology*, Vol. 42, pp. 147–177. [28]
- Land, M.F. and Collett, T.S. (1974). Chasing behaviour of houseflies (*Fannia canicularis*). *Journal of comparative physiology*, Vol. 89, pp. 331–357. [11, 25, 38, 39, 58, 60]
- Laughlin, S.B. (1981). Neural principles in the peripheral visual systems of invertebrates. *Handbook of sensory physiology* (ed. H. Autrum), Vol. VII/6B, pp. 133–280, Springer-Verlag, Berlin. [4, 17, 28, 31, 42, 44, 61, 122]

- Laughlin, S.B. (1996). Matched filtering by a photoreceptor membrane. *Vision research*, Vol. 36, No. 11, pp. 1529–1541. [19, 33, 34, 43, 59, 60, 67]
- Laughlin, S.B. (2001). Energy as a constraint on the coding and processing of sensory information. *Current opinion in neurobiology*, Vol. 11, pp. 475–480. [19, 34]
- Laughlin, S.B. and Weckström, M. (1993). Fast and slow photoreceptors – a comparative-study of the functional diversity of coding and conductances in the Diptera. *Journal of comparative physiology A*, Vol. 172, pp. 593–609. [58]
- Laughlin, S.B., Anderson, J.C., O'Carroll, D., and De Ruyter van Steveninck, R. (2000). Coding efficiency and the metabolic cost of sensory and neural information. *Information theory and the brain* (eds. R. Baddeley, P. Hancock, and P. Földiák), pp. 41–61, Cambridge University Press, Cambridge. [34, 57, 59]
- Lehrer, M. (1997). Active acquisition of depth information by honey-bees. *From living eyes to seeing machines* (eds. M.V. Srinivasan and S. Venkatesh), pp. 37–51, Oxford University Press, Oxford. [27]
- Lewen, G.D., Bialek, W., and De Ruyter van Steveninck, R.R. (2001). Neural coding of naturalistic motion stimuli. *Network: computation in neural systems journal*, Vol. 12, pp. 317–329. [15, 83]
- Lisberger, S.G., Morris, E.J., and Tychsen, L. (1987). Visual motion processing and sensory-motor integration for smooth pursuit eye movements. *Annual review of neuroscience*, Vol. 10, pp. 97–129. [38]
- Livingstone, M. and Hubel, D. (1988). Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science*, Vol. 240, pp. 740–749. [6, 24, 33, 130, 131]
- Marcos, S., Tornow, R-P., Elsner, A.E., and Navarro, R. (1997). Foveal cone spacing and cone photopigment density difference: objective measurements in the same subjects. *Vision research*, Vol. 37, No. 14, pp. 1909–1915. [16]
- Mastebroek, H.A.K., Zaagman, W.H., and Lenting, B.P.M. (1980). Movement detection: performance of a wide-field element in the visual system of the blowfly. *Vision research*, Vol. 20, pp. 467–474. [12, 26, 115, 133]
- Maunsell, J.H.R., Ghose, G.M., Assad, J.A., McAdams, C.J., Boudreau, C.E., and Noerager, B.D. (1999). Visual response latencies of magnocellular and parvocellular LGN neurons in macaque monkeys. *Visual neuroscience*, Vol. 16, No. 1, pp. 1–14. [107]
- McClelland, James L. and Rumelhart, David E. (1991). *Explorations in parallel distributed processing*. MIT Press, Cambridge, MA, 5th edition. [15]

- Meiry, J.L. (1971). Vestibular and proprioceptive stabilization of eye movements. *The control of eye movements* (eds. P. Bach-Y-Rita and C.C. Collins), pp. 483–496, Academic Press, New York, NY. [130]
- Menzel, R. (1979). Spectral sensitivity and color vision in invertebrates. *Handbook of sensory physiology* (ed. H. Autrum), Vol. VII/6A, pp. 503–580, Springer-Verlag, Berlin. [131]
- Miles, F.A. and Wallman, J. (eds.) (1993). *Visual motion and its role in the stabilisation of gaze*, Vol. 5 of *Reviews of oculomotor research*. Elsevier, Amsterdam. [25]
- Miller, W.H. (1979). Ocular optical filtering. *Handbook of sensory physiology* (ed. H. Autrum), Vol. VII/6A, pp. 69–143, Springer-Verlag, Berlin. [2]
- Missler, J.M. and Kamangar, F.A. (1995). A neural network for pursuit tracking inspired by the fly visual system. *Neural networks*, Vol. 8, No. 3, pp. 463–480. [46]
- Mitchell, T. (1997). *Machine learning*. McGraw-Hill, New York, NY. [89]
- Möller, R., Lambrinos, D., Pfeifer, R., Labhart, T., and Wehner, R. (1998). Modeling ant navigation with an autonomous agent. *From animals to animats 5: proceedings of the fifth international conference on simulation of adaptive behavior* (eds. R. Pfeifer, B. Blumberg, J.-A. Meyer, and S.W. Wilson), pp. 185–194, MIT Press, Cambridge, MA. [132]
- Mondada, F., Franzi, E., and Ienne, P. (1994). Mobile robot miniaturisation: a tool for investigation in control algorithms. *Experimental Robotics III* (eds. T. Yoshikawa and F. Miyazaki), Vol. 200 of *Lecture Notes in Control and Information Sciences*, pp. 501–513, Springer-Verlag, Heidelberg. [40]
- Murakami, I. and Cavanagh, P. (1998). A jitter after-effect reveals motion-based stabilization of vision. *Nature*, Vol. 395, pp. 798–801. [61]
- Nakayama, K. (1985). Biological image motion processing: a review. *Vision research*, Vol. 25, No. 5, pp. 625–660. [25]
- Neumann, T.R. and Bülthoff, H. (2001). Insect inspired visual control of translatory flight. *Advances in artificial life* (eds. J. Kelemen and P. Sosik), Vol. 2159 of *Lecture notes in artificial intelligence*, pp. 627–636, Springer-Verlag, Berlin. [65, 116]
- O'Carroll, D.C., Bidwell, N.J., Laughlin, S.B., and Warrant, E.J. (1996). Insect motion detectors matched to visual ecology. *Nature*, Vol. 382, pp. 63–66. [6, 47, 58, 71, 75, 122, 133]

- O'Carroll, D.C., Laughlin, S.B., Bidwell, N.J., and Harris, R.A. (1997). Spatio-temporal properties of motion detectors matched to low image velocities in hovering insects. *Vision research*, Vol. 37, No. 23, pp. 3427–3439. [116, 133]
- Oja, E. (1995). Principal component analysis. *The handbook of brain theory and neural networks* (ed. M.A. Arbib), pp. 753–756, MIT Press, Cambridge, MA. [15]
- Oppenheim, A.V., Willsky, A.S., and Nawab, S.H. (1997). *Signals and systems*. Prentice Hall, London, 2nd edition. [12, 16, 67, 72, 85, 145]
- Pack, Ch., Grossberg, S., and Mingolla, E. (2001). A neural model of smooth pursuit control and motion perception by cortical area MST. *Journal of cognitive neuroscience*, Vol. 13, No. 1, pp. 102–120. [20, 61]
- Peli, E. (1990). Contrast in complex images. *Journal of the optical society of America A*, Vol. 7, pp. 2032–2040. [101]
- Pelli, D.G. (1990). The quantum efficiency of vision. *Vision: coding and efficiency* (ed. C. Blakemore), pp. 3–24, Cambridge University Press, Cambridge. [66, 82]
- Pfeifer, R. and Scheier, C. (1998). Representation in natural and artificial agents: an embodied cognitive science perspective. *Zeitschrift für Naturforschung*, Vol. 53, No. c, pp. 480–503. [19, 20]
- Pfeifer, R. and Scheier, C. (1999). *Understanding intelligence*. MIT-press, Cambridge, MA. [18]
- Pick, B. and Buchner, E. (1979). Visual movement detection under light- and dark-adaptation in the fly, *Musca domestica*. *Journal of comparative physiology A*, Vol. 134, pp. 45–54. [44, 58, 133]
- Poggio, T. and Reichardt, W. (1976). Visual control of orientation behaviour in the fly. II. Towards the underlying neural interactions. *Quarterly reviews of biophysics*, Vol. 9, pp. 377–438. [61, 72]
- Postma, E.O., Van den Herik, H.J., and Hudson, P.T.W. (1996). Robust feedforward processing in synfire chains. *International journal of neural systems*, Vol. 7, No. 4, pp. 537–542. [19]
- Qiu, X., Vanhoutte, K.A.J., Stavenga, D.G., and Arikawa, K. (2002). Ommatidial heterogeneity in the compound eye of the male small white butterfly, *Pieris rapae crucivora*. *Cell tissue research*, Vol. 307, No. 3, pp. 371–379. [14]
- Reichardt, W. (1961). Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. *Sensory communication* (ed. W.A. Rosenblith), pp. 303–317, MIT Press/Wiley, New York, NY. [12, 13, 14, 25, 66]

- Reinecke, R.D. and Cogan, D.G. (1958). Standardization of objective visual acuity measurements: optokinetic nystagmus vs Snellen acuity. *Archives of ophthalmology*, Vol. 60, pp. 418–421. [13]
- Revest, P. and Longstaff, A. (1998). *Molecular neuroscience*. BIOS Scientific Publishers, Oxford. [16]
- Rind, F.C. (1997). Collision avoidance: from the locust eye to a seeing machine. *From living eyes to seeing machines* (eds. M.V. Srinivasan and S. Venkatesh), Oxford University Press, Oxford. [132]
- Rodieck, R.W. (1965). Quantitative analysis of cat retinal ganglion cell response to visual stimuli. *Vision research*, Vol. 5, No. 11, pp. 583–601. [67]
- Ruderman, D.L. (1994). The statistics of natural images. *Network: computation in neural systems journal*, Vol. 5, pp. 517–548. [73]
- Sandini, G., Panerai, F., and Miles, F.A. (2001). The role of inertial and visual mechanisms in the stabilization of gaze in natural and artificial systems. *Motion vision* (eds. J.M. Zanker and J. Zeil), pp. 189–218, Springer-Verlag, Berlin. [65, 66]
- Schilstra, C. and Van Hateren, J.H. (1998). Using miniature sensor coils for simultaneous measurement of orientation and position of small, fast-moving animals. *Journal of neuroscience methods*, Vol. 83, pp. 125–131. [12]
- Schmid, K.L. and Wildsoet, C.F. (1998). Assessment of visual acuity and contrast sensitivity in the chick using an optokinetic nystagmus paradigm. *Vision research*, Vol. 38, pp. 2629–2634. [12]
- Scholes, J.H. (1976). Neuronal connections and cellular arrangement in the fish retina. *Neural principles in vision* (eds. F. Zettler and R. Weiler), pp. 63–93, Springer-Verlag, Berlin. [14]
- Seiffert, A.E. and Cavanagh, P. (1999). Position-based motion perception for color and texture stimuli: effects of contrast and speed. *Vision research*, Vol. 39, pp. 4172–4185. [25]
- Shannon, C.E. and Weaver, W. (1949). *The mathematical theory of communication*. University Press, Urbana, IL. [18]
- Shapley, R. (2000). The receptive fields of visual neurons. *Seeing – handbook of perception and cognition (2nd edition)* (ed. K.K. De Valois), pp. 55–78, Academic Press, San Diego, CA. [14]
- Shepherd, G.M. (ed.) (1990). *The synaptic organization of the brain*. Oxford University Press, New York, NY, 3rd edition. [14]

- Shepherd, G.M. and Koch, Ch. (1990). Introduction to synaptic circuits. *The synaptic organization of the brain* (ed. G.M. Shepherd), pp. 3–31, Oxford University Press, New York, NY. [10]
- Sjöstrand, F.S. (2001). A neural mechanism improving visual acuity by contrast enhancement, a study of foveal vision. *Journal of submicroscopic cytology and pathology*, Vol. 33, Nos. 1–2, pp. 1–5. [30]
- Snippe, H.P. and Koenderink, J.J. (1994). Extraction of optical velocity by use of multi-input Reichardt detectors. *Journal of the optical society of America A*, Vol. 11, No. 4, pp. 1222–1236. [66, 69, 116]
- Snyder, A.W. (1979). The physics of vision in compound eyes. *Handbook of sensory physiology* (ed. H. Autrum), Vol. VII/6A, pp. 225–313, Springer-Verlag, Berlin. [3, 28, 30, 40, 42, 44, 58, 66, 67, 83]
- Sobel, E.C. (1990). The locust's use of motion parallax to measure distance. *Journal of comparative physiology A*, Vol. 167, pp. 579–588. [27]
- Softky, W. and Koch, Ch. (1993). The highly irregular firing of cortical cells is inconsistent with temporal integration of random epsp's. *Journal of neuroscience*, Vol. 13, pp. 334–350. [15]
- Softky, W. and Koch, Ch. (1995). Single-cell models. *The handbook of brain theory and neural networks* (ed. M.A. Arbib), pp. 879–884, MIT Press, Cambridge, MA. [16, 67]
- Srinivasan, M.V. and Bernard, G.D. (1975). The effect of motion on visual acuity of the compound eye: a theoretical analysis. *Vision research*, Vol. 15, pp. 515–525. [12, 31, 42]
- Srinivasan, M.V., Poteser, M., and Kral, K. (1999). Motion detection in insect orientation and navigation. *Vision research*, Vol. 39, No. 16, pp. 2749–2766. [117]
- Stanley, G.B. (2002). Adaptive spatiotemporal receptive field estimation in the visual pathway. *Neural computation*, Vol. 14, pp. 2925–2946. [15]
- Stavenga, D.G. (2002). Colour in the eyes of insects. *Journal of comparative physiology A*, Vol. 188, pp. 337–348. [131]
- Stein, B.E. and Meredith, M.A. (1994). *The merging of the senses*. MIT Press, Cambridge, MA, 2nd edition. [130]
- Strausfeld, N.J. (1976). *Atlas of the insect brain*. Springer-Verlag, Berlin. [14, 33]
- Tatler, B., O'Carroll, D.C., and Laughlin, S.B. (2000). Temperature and the temporal resolving power of fly photoreceptors. *Journal of comparative physiology A*, Vol. 186, pp. 399–407. [4]

- Thibos, L.N. (2000). Retinal image formation and sampling. *Seeing - handbook of perception and cognition (2nd edition)* (ed. K.K. De Valois), pp. 1–54, Academic Press, San Diego, CA.[61]
- Tinbergen, N. (1951). *The study of instinct*. Oxford University Press, New York, NY.[11]
- Tolhurst, D. J. and Tadmor, Y. (1997). Band-limited contrast in natural images explains the detectability of changes in the amplitude spectra. *Vision research*, Vol. 37, No. 23, pp. 3203–3215.[73, 101]
- Toveé, M.J. (1996). *An introduction to the visual system*. Cambridge university press, Cambridge.[4]
- Ullman, S. (1983). The measurement of visual motion — computational considerations and some neurophysiological implications. *Trends in neuroscience*, Vol. 6, pp. 177–179.[25]
- Van der Schaaf, A. and Van Hateren, J.H. (1996). Modelling the power spectra of natural images: statistics and information. *Vision research*, Vol. 36, No. 17, pp. 2759–2770.[73, 102]
- Van Essen, D.C. and Anderson, C.H. (1990). Information processing strategies and pathways in the primate retina and visual cortex. *An introduction to neural and electronic networks* (eds. S.F. Zorner, J.L. Davis, and C. Lau), pp. 43–72, Academic Press, San Diego, CA.[33]
- Van Essen, D.C., Anderson, Ch.H., and Felleman, D.J. (1992). Information processing in the primate visual system: an integrated systems perspective. *Science*, Vol. 255, pp. 419–423.[12]
- Van Hateren, J.H. (1992a). Real and optimal neural images in early vision. *Nature*, Vol. 360, pp. 68–70.[15]
- Van Hateren, J.H. (1992b). Theoretical predictions of spatiotemporal receptive fields of fly LMCs, and experimental validation. *Journal of comparative physiology A*, Vol. 171, pp. 157–170.[19, 66, 116]
- Van Hateren, J.H. (1993). Spatiotemporal contrast sensitivity of early vision. *Vision research*, Vol. 33, No. 2, pp. 257–267.[73, 82, 83, 101, 116]
- Van Hateren, J.H. (1997). Processing of natural time series of intensities by the visual system of the blowfly. *Vision research*, Vol. 37, No. 23, pp. 3407–3416.[73]
- Van Hateren, J.H. and Snippe, H.P. (2001). Information theoretical evaluation of parametric models of gain control in blowfly photoreceptor cells. *Vision research*, Vol. 41, pp. 1851–1865.[12]

- Van Hateren, J. H. and Van der Schaaf, A. (1998). Independent component filters of natural images compared with simple cells in primary visual cortex. *Proceeding of the Royal Society of London B*, Vol. 265, pp. 359–366. [101, 102]
- Van Santen, J.P.H. and Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the optical society of America A*, Vol. 2, No. 2, pp. 300–321. [25, 70]
- Varela, F.J., Thompson, E., and Rosch, E. (1991). *The embodied mind*. MIT Press, Cambridge, MA. [20]
- Viollett, S. and Franceschini, N. (1999). Biologically-inspired visual scanning sensor for stabilization and tracking. *IROS'99 Conference on Intelligent Robots and Systems*, Vol. 1, pp. 204–209, Kyongju. [20]
- Wandell, B.A. (1995). *Foundations of vision*. Sinauer Associates, Sunderland, MA. [4, 27, 30, 39, 98, 103, 104, 131]
- Warrant, E.J. (1999). Seeing better at night: life style, eye design and the optimum strategy of spatial and temporal summation. *Vision research*, Vol. 39, pp. 1611–1630. [2, 3, 4, 6, 27, 28, 30, 31, 42, 58, 83, 94, 115, 133, 136]
- Warzecha, A-K. and Egelhaaf, M. (1996). Intrinsic properties of biological motion detectors prevent the optomotor control system from getting unstable. *Philosophical transactions of the Royal Society of London B*, Vol. 351, pp. 1579–1591. [65, 66, 83, 86, 87, 94]
- Warzecha, A-K. and Egelhaaf, M. (2000). Response latency of a motion-sensitive neuron in the fly visual system: dependence on stimulus parameters and physiological conditions. *Vision research*, Vol. 40, pp. 2973–2983. [108]
- Webb, B. (2001). Can robots make good models of biological behaviour? *The behavioral and brain sciences*, Vol. 24, No. 6, pp. 1033–1049. [18]
- Webb, B. (2002). Robots in invertebrate neuroscience. *Nature*, Vol. 417, pp. 359–363. [19, 20, 144]
- Weckström, M., Hardie, R.C., and Laughlin, S.B. (1991). Voltage-activated potassium channels in blowfly photoreceptors and their role in light adaptation. *Journal of physiology*, Vol. 440, pp. 635–657. [16]
- Wehner, R. (1981). Spatial vision in arthropods. *Handbook of sensory physiology* (ed. H. Autrum), Vol. VII/6C, pp. 287–616, Springer-Verlag, Berlin. [4, 24, 25, 37, 43]
- Wehner, R. (1997). Insect navigation: low-level solutions to high-level tasks. *From living eyes to seeing machines* (eds. M.V. Srinivasan and S. Venkatesh), pp. 158–173, Oxford University Press, Oxford. [131]

- Wehrhahn, C., Poggio, T., and Bülthoff, H. (1982). Tracking and chasing in houseflies (*Musca*). *Biological cybernetics*, Vol. 45, pp. 123–130. [61]
- Wiersma, C.A.G. and Yamaguchi, T. (1967). Integration of visual stimuli by the crayfish central nervous system. *Journal of experimental biology*, Vol. 47, pp. 409–431. [25]
- Yablonskiy, D.A., Ackerman, J.J.H., and Raichle, M.E. (2000). Coupling between changes in human brain temperature and oxidative metabolism during prolonged visual stimulation. *Proceedings of the National Academy of Sciences*, Vol. 97, No. 13, pp. 7603–7608. [33]
- Young, L.R. (1971). Pursuit eye tracking movements. *The control of eye movements* (eds. P. Bach-y-Rita, C.C. Collins, and J.E. Hyde), pp. 429–443, Academic Press, New York, NY. [58]
- Zaagman, W.H., Mastebroek, H.A.K., and De Ruyter van Steveninck, R. (1983). Adaptive strategies in fly vision: on their image-processing qualities. *IEEE transactions on systems, man, and cybernetics*, Vol. 13, No. 5, pp. 900–906. [33]
- Zanker, J.M. and Zeil, J. (2001a). Processing motion in the real world. *Motion vision* (eds. J.M. Zanker and J. Zeil), Springer-Verlag, Berlin. [14, 20]
- Zanker, J.M. and Zeil, J. (eds.) (2001b). *Motion vision*. Springer-Verlag, Berlin. [25, 63]
- Zanker, J.M., Srinivasan, M.V., and Egelhaaf, M. (1999). Speed tuning in elementary motion detectors of the correlation type. *Biological cybernetics*, Vol. 80, pp. 109–116. [69]
- Zenker, W. and Kubik, S. (1996). Brain cooling in humans – anatomical considerations. *Anatomy and embryology*, Vol. 193, pp. 1–13. [32]

Summary

Animals act in a visual world. To perceive the visual world accurately, many animals have developed dedicated mechanisms to detect fast changes. For instance, the light-sensitive cells in the eyes of the blowfly *Calliphora* are well able to follow the switching on and off of TL-tubes (100 times per second). Since people detect changes much less well than blowflies do, they perceive a TL-tube as a constant light source. The ability to detect changes is termed 'temporal resolution'. The example above shows that the temporal resolution has developed much better in the eyes of blowflies than in those of humans.

In contrast to the high temporal resolution it is generally known that blowflies detect little detail in still images. For instance, from earlier research it has appeared that blowflies distinguish much less detail than humans do. This difference is caused by the design of the eyes of humans and blowflies. The level of detail that an animal may detect is called 'spatial resolution'. Therefore, people possess a higher spatial resolution than blowflies do. The highest spatial resolution is found in the eyes of birds of prey that need to detect small prey species on the surface from high altitudes. The research in this thesis aims to explain why some animals have developed the ability to see much detail in still images (high spatial resolution; *e.g.*, humans and birds of prey) while others have developed a visual system that detects fast changes (high temporal resolution; *e.g.*, blowflies).

Maximising both spatial and temporal resolution is important for accurate perception of the visual world. However, from the neurobiological literature it appears that the simultaneous maximisation of both types of resolution does not occur in nature. It is even suggested that a high spatial resolution is only found when the temporal resolution is low, and *vice versa*. It is generally assumed that such a trade-off between spatial and temporal resolution arises to optimise the visually guided behaviour of animals to their ecological habitats. For instance, the habitat of birds of prey contains small prey species. Therefore, the visually guided hunting behaviour of such birds is optimised when they are able to detect much detail in visual images (when they possess a high spatial resolution as was indicated above). Up to now it has been undecided which are the generic causes of the trade-off between spatial and temporal resolution. Generic causes are defined as factors that (a) explain the existence of a trade-off in many occurrences of visually guided behaviour, that (b) cause a trade-off in biological visual systems, and that (c) cause a trade-off in

artificial visual systems. Therefore, the central question in the research, formulated in chapter 1, reads: "which are the generic causes of the trade-off between spatial and temporal resolution in visually guided behaviour?"

Answering this question starts with selecting an appropriate research approach. In chapter 2, a taxonomy is established to categorise the different approaches to studying visually guided behaviour. After careful comparison of these approaches, it is decided to use so called situated computer models. These models simulate the visually guided behaviour of animals in their ecological habitat. Several arguments indicate that this research approach is more appropriate than its two main alternatives: controlled behavioural experiments on animals and the use of information theory.

Chapter 3 investigates how the research approach may be optimally used to answer the central research question. To this aim the two most important control systems of visually guided behaviour in animals are selected: position-detection systems and motion-detection systems. These two biological control systems serve as a starting point for the construction of situated models of visually guided behaviour. Subsequently, five so-called determinants are identified as candidate causes of the trade-off: light sensitivity, contrast sensitivity, motion sensitivity, heat production, and energy consumption. These determinants are properties of visual systems that (1) are important for visually guided behaviour, that (2) depend on the spatial and temporal resolution, and that (3) may cause a trade-off. Finally, a procedure is established to determine whether the determinants constitute the generic causes of the trade-off. The procedure involves the maximisation of the behavioural success of the situated models of visually guided behaviour that were mentioned above. The success is maximised by tuning the spatial and temporal resolution to values that optimise the determinants (optimal light sensitivity, optimal contrast sensitivity, etc.). The maximisation process is repeated for different simulated ecological habitats. If the tuned resolution values are positioned on a so-called trade-off curve, it is concluded that the determinants cause a trade-off in the models that are used. A trade-off curve describes how the spatial resolution is traded for temporal resolution, and *vice versa*. The procedure concludes with a comparison between the models and a set of biological systems to verify that the causes are generic.

In chapter 4 the procedure is applied to a situated model of position-detection systems. An evolutionary algorithm maximises the behavioural success of the model by optimising the values of spatial and temporal resolution. The optimal resolution values are positioned on a trade-off curve. This observation leads to the conclusion that the determinants cause a trade-off between spatial and temporal resolution in the model. Moreover, it follows from the experiment that the velocity of moving targets determines whether the model develops a high temporal or a high spatial resolution. This result provides an explanation for the differences in resolution values of humans and blowflies that were mentioned above. The velocity of moving targets in an ecological habitat determines the trade-off between spatial and temporal resolution. In the daily activities of blowflies the average velocity of targets is much higher than in the case of humans. Therefore, it is likely that the differences

between the ecological habitats of blowflies and humans determine that the former species has developed a high temporal resolution, whereas the visual system of the latter species has specialised in a high spatial resolution.

In chapter 5 the trade-off is studied in a model of motion-detection systems. To this aim the behavioural success of the model is maximised by setting the resolution values manually. The experiment differs from the one in chapter 4 in two important ways. First, the detection of position differs fundamentally from the detection of motion. Second, the experiment in chapter 5 takes into consideration how the statistical properties of natural visual images influence the control of behaviour. The results generalise the conclusions of chapter 4 on position-detection systems: the determinants cause a trade-off in the model of motion-detection systems as well.

Arguments to choose situated modelling as our research approach have been provided in chapter 2. The appropriateness of the approach is verified later in chapter 6. For this, the results from chapters 4 and 5 are compared to those of the two main alternative approaches. The comparison leads to the following two findings. First, in qualitative comparison to controlled behavioural experiments on animals the situated modelling studies yield results that (1) are better comparable to each other, that (2) are available in unlimited quantities, and that (3) lead to a better understanding of the trade-off. Second, in quantitative comparison to the use of information theory it is determined that situated modelling produces results that are more plausible from an evolutionary point of view. In conclusion it is stated that the use of situated modelling outperforms the two main alternative approaches when applied to finding the generic causes of the trade-off.

The results from chapters 4 and 5 show that the five determinants cause a trade-off between spatial and temporal resolution in the two models that are employed. In chapter 7 it is determined whether the causes are generic. To this purpose, two critical notes are raised: first, besides the five original determinants of the trade-off, the experiments mentioned above have yielded ten additional determinants. Therefore, the original determinants may not uniquely constitute the generic causes of the trade-off. Second, additional studies have revealed that the two selected systems (that detect position or motion in visual scenes) do not control all occurrences of visually guided behaviour. Therefore, the determinants may not be truly generic. To resolve the critical notes, the resolution of the situated models is compared to that of animals. To this aim, a set of biological resolution values is compiled. After matching the resolution of models and animals it has been found that the five original determinants explain a significant part of the ordering of the biological resolution values. The remaining part is explained by taking into consideration three additional determinants. Besides a comparison to animals, the chapter shows that the five determinants cause a trade-off in artificial visual systems as well. Based on the two comparisons, it is concluded that the five original determinants are the most important causes of the trade-off. Future research is expected to reveal what influence the additional determinants and the alternative occurrences of visually guided behaviour have on the composition of the generic causes of the trade-off.

Chapter 8 answers the research question as follows. The five determinants are

the generic causes of the trade-off between spatial and temporal resolution for the following three reasons. First, they cause a trade-off in many occurrences of visually guided behaviour. Second, they explain the differences between the resolution values of different animals (such as humans and blowflies). Third, they yield principles that may be used to determine the optimal resolution of artificial visual systems. Aside from the aforementioned conclusions, the chapter concludes that the research approach used (situated modelling) is important for studying the trade-off between spatial and temporal resolution as it outperforms its two main alternative approaches.

The differences between the spatial and temporal resolution of humans and blowflies, as were illustrated in the beginning of this summary, have been fully explained by the research presented in this thesis. Owing to a number of important properties of visual systems, such as light sensitivity and heat production, and due to the development of high spatial resolution, natural selection voted for a temporal resolution in the human visual system that is lower than in the visual system of blowflies. As for many natural phenomena, the saying “you cannot have your cake and eat it” goes for spatial and temporal resolution as well.

Samenvatting

Dieren bewegen zich in een visuele wereld. Om deze wereld nauwkeurig te kunnen waarnemen hebben veel dieren zich gespecialiseerd in het zien van snelle veranderingen. Zo kunnen de lichtgevoelige cellen in de ogen van de bromvlieg *Calliphora* het aan- en uitgaan van een TL-buis (100 keer per seconde) goed volgen. Omdat mensen veranderingen veel minder goed zien dan bromvliegen, nemen zij een TL-buis waar als een constante lichtbron. Het vermogen om veranderingen waar te nemen wordt aangeduid met de term 'temporele resolutie' en is bij bromvliegen dus veel beter ontwikkeld dan bij mensen.

Tegenover de hoge temporele resolutie staat dat bromvliegen weinig details kunnen waarnemen in een stilstaand beeld. Zo is uit eerder onderzoek gebleken dat zij veel minder details detecteren dan mensen. Dit wordt veroorzaakt door het soort ogen van de bromvlieg. De hoeveelheid detail die een dier kan waarnemen heet 'spatiële resolutie'. Mensen bezitten dus een hogere spatiële resolutie dan bromvliegen. De hoogste spatiële resolutie is gevonden bij roofvogels die vanaf grote hoogten hun prooi op de grond zoeken. Het onderzoek in dit proefschrift beoogt uit te leggen waarom sommige dieren zich gespecialiseerd hebben in het zien van details in stilstaande beelden (hoge spatiële resolutie; voorbeelden zijn mensen en roofvogels) en andere in het zien van snelle veranderingen (hoge temporele resolutie; een voorbeeld is de bromvlieg).

Het maximaliseren van zowel de spatiële als de temporele resolutie is belangrijk voor een nauwkeurige visuele waarneming. Uit de neurobiologische literatuur blijkt echter dat de gelijktijdige maximalisatie van beide soorten resolutie niet voorkomt bij dieren. Men vermoedt zelfs dat een hoge spatiële resolutie alleen voorkomt wanneer de temporele resolutie laag is, en omgekeerd. Deze uitwisseling van resolutiewaarden wordt aangeduid met de Engelse term 'trade-off'. Aangenomen wordt dat de trade-off tussen spatiële en temporele resolutie optreedt om het visueel gestuurd gedrag van dieren te optimaliseren voor hun ecologische leefomgeving. Zo bevat de ecologische leefomgeving van roofvogels kleine prooidieren. Het visueel gestuurd jaaggedrag van een roofvogel is daarom optimaal als hij veel details kan zien (hoge spatiële resolutie, zie hierboven). Tot nu toe was het evenwel onbekend welke factoren de generieke oorzaken zijn van de trade-off. Generieke oorzaken zijn de factoren die (a) de trade-off verklaren in veel soorten visueel gestuurd gedrag, die (b) een trade-off veroorzaken in biologische visuele systemen en die (c) een trade-off veroorzaken

in kunstmatige visuele systemen. De centrale vraag in het onderzoek (hoofdstuk 1) luidt derhalve: "Welke factoren zijn de generieke oorzaken van de trade-off tussen spatiële en temporele resolutie in visueel gestuurd gedrag?"

Het beantwoorden van deze vraag begint met het selecteren van een onderzoeksbenadering. In hoofdstuk 2 wordt een taxonomie opgesteld om de verschillende benaderingen, die visueel gestuurd gedrag onderzoeken, te categoriseren. Aan de hand van een zorgvuldige vergelijking wordt besloten zogenaamde gesitueerde computermodellen te gebruiken. Gesitueerde computermodellen simuleren het visueel gestuurd gedrag van dieren in hun ecologische leefomgeving. Er wordt met argumenten aangegeven dat deze benadering geschikter is dan de twee belangrijkste alternatieven: gecontroleerde gedragsexperimenten met dieren en het gebruik van informatietheorie.

In hoofdstuk 3 wordt onderzocht hoe de gekozen benadering kan worden ingezet om de onderzoeksvraag te beantwoorden. Daartoe worden eerst de twee belangrijkste controle systemen voor visueel gestuurd gedrag in dieren geselecteerd: positiedetectiesystemen en bewegingsdetectiesystemen. Deze twee biologische systemen dienen als uitgangspunt voor gesitueerde modellen van visueel gestuurd gedrag. Vervolgens worden vijf zogeheten determinanten geïdentificeerd als kandidaat-oorzaken van de trade-off: lichtgevoeligheid, contrastgevoeligheid, bewegingsgevoeligheid, warmteproductie en energieverbruik. Determinanten zijn eigenschappen van visuele systemen die (1) belangrijk zijn voor visueel gestuurd gedrag, (2) afhankelijk zijn van de spatiële en temporele resolutie en (3) een trade-off kunnen veroorzaken. Ten slotte wordt een procedure opgesteld om te bepalen of de determinanten de generieke oorzaken van de trade-off zijn. In deze procedure wordt het gedragsmatig succes van bovengenoemde modellen gemaximaliseerd. Dit gebeurt door de spatiële en temporele resolutie in te stellen op waarden die corresponderen met optimale determinanten (dat wil zeggen optimale licht-, contrast- en bewegingsgevoeligheid, etc.). Deze handelingen worden herhaald voor verschillende gesimuleerde ecologische leefomgevingen. Indien de ingestelde waarden van spatiële en temporele resolutie op een zogenaamde trade-off curve liggen volgt de conclusie dat de vijf determinanten de oorzaken van de trade-off in de modellen zijn. Een trade-off curve beschrijft hoe de spatiële resolutie wordt uitgewisseld voor temporele resolutie, en omgekeerd. De procedure besluit met een vergelijking tussen de modellen en een aantal biologische visuele systemen om uit te wijzen of de gevonden oorzaken universeel zijn.

In hoofdstuk 4 wordt de hierboven beschreven procedure toegepast op een model van positiedetectiesystemen. Een evolutionair algoritme maximaliseert het gedragsmatig succes van het model door de waarden van de spatiële en temporele resolutie te optimaliseren. De optimale resolutiewaarden liggen op een trade-off curve. Daaruit wordt geconcludeerd dat de determinanten een trade-off tussen de spatiële en temporele resolutie veroorzaken in het model. Tevens blijkt uit het experiment dat de snelheid van bewegende objecten bepaalt of het model zich specialiseert in de spatiële resolutie, dan wel de temporele resolutie. Dit levert een verklaring op voor de hierboven geobserveerde verschillen in resolutie van de mens en de bromvlieg. De snelheid van objecten in de leefomgeving bepaalt de trade-off tussen de spatiële

en temporele resolutie. Voor bromvliegen is de gemiddelde snelheid van objecten hoger dan voor mensen. Daarom is het aannemelijk dat de verschillende ecologische leefomgevingen bepalen dat bromvliegen zich hebben gespecialiseerd in hoge temporele resolutie, terwijl mensen zich hebben gespecialiseerd in hoge spatiële resolutie.

In hoofdstuk 5 wordt de trade-off onderzocht in een model van bewegingsdetectiesystemen. Hiertoe wordt het gedragsmatige succes van het model gemaximaliseerd. Dit experiment verschilt van dat in hoofdstuk 4 op twee manieren: ten eerste verschilt het detecteren van beweging fundamenteel van het detecteren van positie en ten tweede wordt rekening gehouden met de statistische kenmerken van de natuurlijke leefomgeving van dieren. De resultaten generaliseren de conclusies van hoofdstuk 4 over positiedetectiesystemen naar bewegingsdetectiesystemen: de determinanten veroorzaken een trade-off tussen de spatiële en temporele resolutie van bewegingsdetectiesystemen.

De in hoofdstuk 2 beargumenteerde geschiktheid van de gekozen onderzoeksbenadering (het gebruik van gesitueerde modellen) wordt geverifieerd in hoofdstuk 6. Hiervoor worden de resultaten uit hoofdstukken 4 en 5 vergeleken met die van de belangrijkste alternatieve benaderingen. Dit leidt tot de volgende twee bevindingen. Ten eerste, in kwalitatieve vergelijking tot gecontroleerde gedragsexperimenten met dieren leveren de modelleringstudies resultaten op die (1) beter vergelijkbaar zijn, (2) in grotere getalen beschikbaar zijn en (3) tot een beter begrip van de trade-off leiden. Ten tweede, in kwantitatieve vergelijking tot informatietheorie blijkt dat het gebruik van gesitueerde modellen resultaten oplevert die evolutionair plausibeler zijn. De conclusie luidt dat in dit onderzoek het gebruik van gesitueerde modellen de alternatieve benaderingen overtreft.

De resultaten uit hoofdstukken 4 en 5 tonen aan dat de vijf determinanten een trade-off van resolutiewaarden veroorzaken in de twee gebruikte modellen van visueel gestuurd gedrag. In hoofdstuk 7 wordt bepaald of deze oorzaken universeel zijn. Hiertoe worden eerst twee punten van kritiek geformuleerd. Ten eerste worden tien additionele determinanten opgesomd die werden gevonden in hoofdstukken 4 en 5. Deze opsomming toont aan dat de vijf originele determinanten mogelijk niet de enige oorzaken van de trade-off zijn. Ten tweede worden alternatieve voorkomens van visueel gestuurd gedrag nader bestudeerd. Hieruit blijkt dat de twee geselecteerde systemen (positie- en bewegingsdetectiesystemen) niet alle voorkomens van visueel gestuurd gedrag besturen. Om deze twee punten van kritiek weg te nemen wordt de resolutie van de twee modellen vergeleken met die van dieren. Hiertoe wordt een verzameling van biologische resolutiewaarden samengesteld. De vergelijking toont aan dat de vijf oorspronkelijke determinanten de ordening van deze waarden voor een significant gedeelte verklaren. Het overige gedeelte wordt verklaard wanneer drie additionele determinanten in aanmerking worden genomen. Naast een vergelijking met dieren wordt bepaald dat de vijf determinanten tevens de trade-off in kunstmatige visuele systemen veroorzaken. Op grond van de twee vergelijkingen wordt geconcludeerd dat de vijf originele determinanten de belangrijkste oorzaken van de trade-off zijn. Toekomstig onderzoek moet uitwijzen welke invloed de additionele determinanten en de alternatieve voorkomens van gedrag hebben op de generieke oorzaken van de

trade-off.

Hoofdstuk 8 beantwoordt de onderzoeksvraag als volgt. De vijf determinanten zijn generieke oorzaken van de trade-off om de volgende drie redenen. Ten eerste veroorzaken zij de trade-off tussen spatiële en temporele resolutie in veel verschillende soorten visueel gestuurd gedrag. Ten tweede verklaren zij de verschillen in spatiële en temporele resolutie in verschillende diersoorten (zoals de mens en de bromvlieg). Ten derde leveren zij principes op voor het bepalen van de optimale resolutie van kunstmatige visuele systemen. Daarnaast leidt dit laatste hoofdstuk tot de conclusie dat de gebruikte onderzoeksbenadering (gesitueerd modelleren) belangrijk is voor het onderzoek naar de trade-off tussen spatiële en temporele resolutie.

De verschillen in de visuele capaciteiten van de mens en de bromvlieg, zoals die zijn geschetst in het begin van deze samenvatting, worden verklaard door het onderzoek in dit proefschrift. Vanwege een aantal belangrijke eigenschappen van visuele systemen, zoals lichtgevoeligheid en warmteproductie, en vanwege hun hoge spatiële resolutie, hebben mensen een lagere temporele resolutie dan bromvliegen. Zoals voor veel natuurlijke fenomenen blijkt ook voor spatiële en temporele resolutie te gelden: het is kiezen of delen.

Curriculum Vitae

Laurens Jan (Rens) Kortmann was born in Rotterdam, the Netherlands, on 8 October 1975. He attended secondary school, Lorentz Lyceum, in Eindhoven from 1987 until 1993 and received the diploma *Gymnasium β* . He started his undergraduate studies at the University of Groningen reading applied physics in 1993 and continued his studies in 1994 in the programme of cognitive science and engineering. He was a student assistant in 1996 and 1997. His M.Sc. degree was received in 1998 by carrying out a research project at the University of Edinburgh, Scotland. There he studied a robot model of mate-finding behaviour in crickets. After graduating he started to work as a research assistant (AIO) at the University of Maastricht in 1998. At the Institute for Knowledge and Agent Technology (IKAT) he investigated computer models of visually guided behaviour in animals under the supervision of professors H.J. van den Herik and E.O. Postma until 2002. The research resulted in several publications and this thesis. Besides performing scientific tasks, he was engaged in co-founding a robotics laboratory, in teaching (knowledge engineering and psychology), in organising conferences and workshops, and in coordinating a colloquium series.

SIKS dissertation series

1998

- 1998-1 Johan van den Akker (CWI) DEGAS - An Active, Temporal Database of Autonomous Objects
- 1998-2 Floris Wiesman (UM) Information Retrieval by Graphically Browsing Meta-Information
- 1998-3 Ans Steuten (TUD) A Contribution to the Linguistic Analysis of Business Conversations within the Language/Action Perspective
- 1998-4 Dennis Breuker (UM) Memory versus Search in Games
- 1998-5 E.W.Oskamp (RUL) Computerondersteuning bij Straftoemeting

1999

- 1999-1 Mark Sloof (VU) Physiology of Quality Change Modelling; Automated modelling of Quality Change of Agricultural Products
- 1999-2 Rob Potharst (EUR) Classification using decision trees and neural nets
- 1999-3 Don Beal (UM) The Nature of Minimax Search
- 1999-4 Jacques Penders (UM) The practical Art of Moving Physical Objects
- 1999-5 Aldo de Moor (KUB) Empowering Communities: A Method for the Legitimate User-Driven Specification of Network Information Systems
- 1999-6 Niek J.E. Wijngaards (VU) Re-design of compositional systems
- 1999-7 David Spelt (UT) Verification support for object database design
- 1999-8 Jacques H.J. Lenting (UM) Informed Gambling: Conception and Analysis of a Multi-Agent Mechanism for Discrete Reallocation

2000

- 2000-1 Frank Niessink (VU) Perspectives on Improving Software Maintenance
- 2000-2 Koen Holtman (TUE) Prototyping of CMS Storage Management

- 2000-3 Carolien M.T. Metselaar (UvA) Sociaal-organisatorische gevolgen van kennistechnologie; een procesbenadering en actorperspectief
- 2000-4 Geert de Haan (VU) ETAG, A Formal Model of Competence Knowledge for User Interface Design
- 2000-5 Ruud van der Pol (UM) Knowledge-based Query Formulation in Information Retrieval
- 2000-6 Rogier van Eijk (UU) Programming Languages for Agent Communication
- 2000-7 Niels Peek (UU) Decision-theoretic Planning of Clinical Patient Management
- 2000-8 Veerle Coupé (EUR) Sensitivity Analysis of Decision-Theoretic Networks
- 2000-9 Florian Waas (CWI) Principles of Probabilistic Query Optimization
- 2000-10 Niels Nes (CWI) Image Database Management System Design Considerations, Algorithms and Architecture
- 2000-11 Jonas Karlsson (CWI) Scalable Distributed Data Structures for Database Management

2001

- 2001-1 Silja Renooij (UU) Qualitative Approaches to Quantifying Probabilistic Networks
- 2001-2 Koen Hindriks (UU) Agent Programming Languages: Programming with Mental Models
- 2001-3 Maarten van Someren (UvA) Learning as problem solving
- 2001-4 Evgueni Smirnov (UM) Conjunctive and Disjunctive Version Spaces with Instance-Based Boundary Sets
- 2001-5 Jacco van Ossenbruggen (VU) Processing Structured Hypermedia: A Matter of Style
- 2001-6 Martijn van Welie (VU) Task-based User Interface Design
- 2001-7 Bastiaan Schonhage (VU) Diva: Architectural Perspectives on Information Visualization
- 2001-8 Pascal van Eck (VU) A Compositional Semantic Structure for Multi-Agent Systems Dynamics
- 2001-9 Pieter Jan 't Hoen (RUL) Towards Distributed Development of Large Object-Oriented Models, Views of Packages as Classes
- 2001-10 Maarten Sierhuis (UvA) Modeling and Simulating Work Practice BRAHMS: a multiagent modeling and simulation language for work practice analysis and design
- 2001-11 Tom M. van Engers (VU) Knowledge Management: The Role of Mental Models in Business Systems Design

2002

- 2002-01 Nico Lassing (VU) Architecture-Level Modifiability Analysis
- 2002-02 Roelof van Zwol (UT) Modelling and searching web-based document collections
- 2002-03 Henk Ernst Blok (UT) Database Optimization Aspects for Information Retrieval
- 2002-04 Juan Roberto Castelo Valdueza (UU) The Discrete Acyclic Digraph Markov Model in Data Mining
- 2002-05 Radu Serban (VU) The Private Cyberspace Modeling Electronic Environments inhabited by Privacy-concerned Agents
- 2002-06 Laurens Mommers (UL) Applied legal epistemology; Building a knowledge-based ontology of the legal domain
- 2002-07 Peter Boncz (CWI) Monet: A Next-Generation DBMS Kernel For Query-Intensive Applications
- 2002-08 Jaap Gordijn (VU) Value Based Requirements Engineering: Exploring Innovative E-Commerce Ideas
- 2002-09 Willem-Jan van den Heuvel (KUB) Integrating Modern Business Applications with Objectified Legacy Systems
- 2002-10 Brian Sheppard (UM) Towards Perfect Play of Scrabble
- 2002-11 Wouter C.A. Wijngaards (VU) Agent Based Modelling of Dynamics: Biological and Organisational Applications
- 2002-12 Albrecht Schmidt (UvA) Processing XML in Database Systems
- 2002-13 Hongjing Wu (TUE) A Reference Architecture for Adaptive Hypermedia Applications
- 2002-14 Wieke de Vries (UU) Agent Interaction: Abstract Approaches to Modelling, Programming and Verifying Multi-Agent Systems
- 2002-15 Rik Eshuis (UT) Semantics and Verification of UML Activity Diagrams for Workflow Modelling
- 2002-16 Pieter van Langen (VU) The Anatomy of Design: Foundations, Models and Applications
- 2002-17 Stefan Manegold (UvA) Understanding, Modeling, and Improving Main-Memory Database Performance

2003

- 2003-01 Heiner Stuckenschmidt (VU) Ontology-Based Information Sharing in Weakly Structured Environments
- 2003-02 Jan Broersen (VU) Modal Action Logics for Reasoning About Reactive Systems
- 2003-03 Martijn Schuemie (TUD) Human-Computer Interaction and Presence in Virtual Reality Exposure Therapy

- 2003-04 Milan Petkovic (UT) Content-Based Video Retrieval Supported by Database Technology
- 2003-05 Jos Lehmann (UvA) Causation in Artificial Intelligence and Law - A modelling approach
- 2003-06 Boris van Schooten (UT) Development and specification of virtual environments
- 2003-07 Machiel Jansen (UvA) Formal Explorations of Knowledge Intensive Tasks
- 2003-08 Yong-Ping Ran (UM) Repair-Based Scheduling
- 2003-09 Rens Kortmann (UM) The resolution of visually guided behaviour