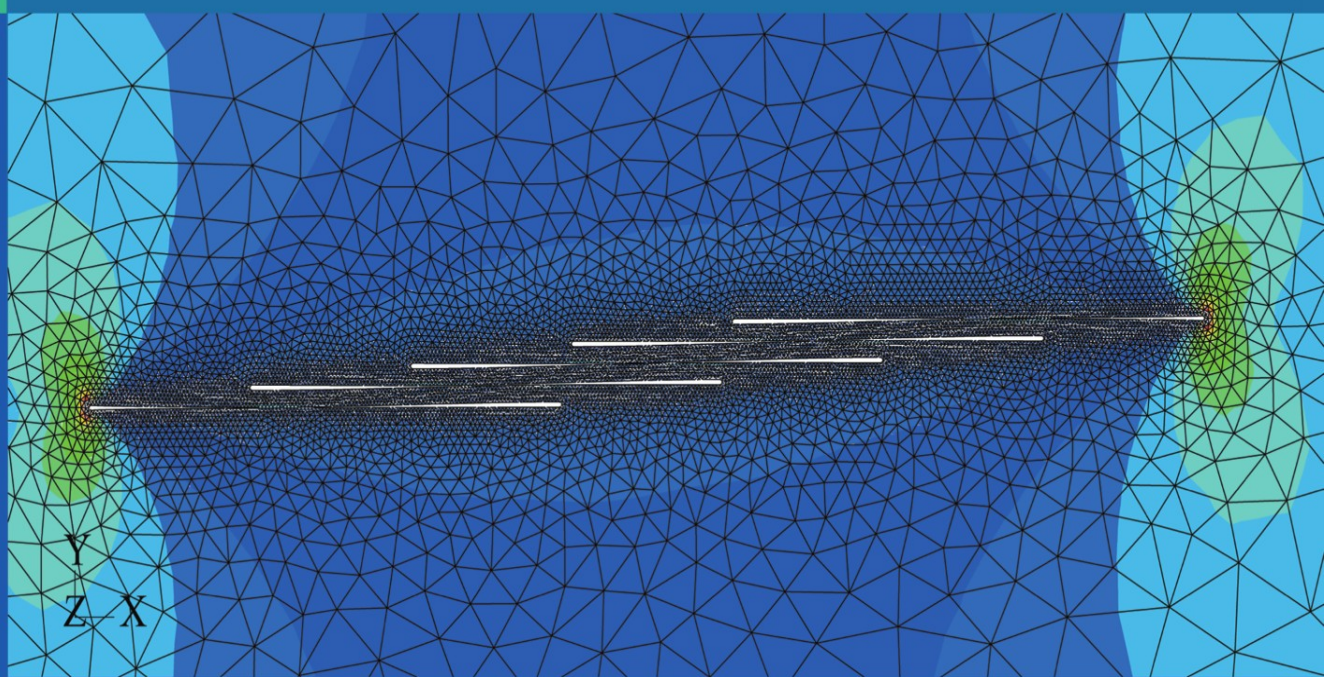
Friedrich G. Barth · Joseph A. C. Humphrey
Mandyam V. Srinivasan *Editors*

Frontiers in Sensing

From Biology to Engineering



SpringerWienNewYork

SpringerWienNewYork

Cover picture: Finite element analysis showing stresses around an array of five slits modelled to study the effect of loads on the deformation of highly sensitive strain sensors in the cuticular skeleton of spiders.

Friedrich G. Barth
Joseph A. C. Humphrey†
Mandyam V. Srinivasan

Frontiers in Sensing

From Biology to Engineering

Friedrich G. Barth
Department of Neurobiology, Faculty of Life Sciences,
University of Vienna, Vienna, Austria

Joseph A. C. Humphrey†
Department of Mechanical and Aerospace Engineering,
University of Virginia, Charlottesville, VA, USA

Mandyam V. Srinivasan
Queensland Brain Institute, University of Queensland
St. Lucia, QLD, Australia

SpringerWienNewYork is part of
Springer Science+Business Media
springer.at

Cover Illustrations: © B. Hößl, H. Böhm, F. G. Rammerstorfer, F. G. Barth/
Vienna University of Technology and University of Vienna
Typesetting: Jung Crossmedia Publishing GmbH, 35633 Lahnau, Germany

Printed on acid-free and chlorine-free bleached paper
SPIN: 12636488

With 186 (mostly coloured) Figures.

Library of Congress Control Number: 2011936879

ISBN 978-3-211-99748-2 SpringerWienNewYork

Preface

Although the idea of *applying principles found in nature* to the design of engineering systems has a long history, only recently has biomimetics begun to establish itself with rapidly increasing impact in both biology and engineering. One of the reasons for this development is a number of highly sophisticated technologies and analytical methods that have now become available, highlighted among others by molecular biology on the one hand and micro- and nano-fabrication technologies in engineering on the other. The general interest in looking upon nature as a database for possible solutions of complex problems that have already been worked out by biological systems seems to have reached a critical threshold from where it is rapidly expanding.

There is *no life without sensors and sensing*. Even at the level of bacteria sensory performance is already well developed, guiding the cell's activity in its species specific environment. Among the higher animals, hundreds of millions of years of evolution and the effect of permanent 'quality control' by selective pressures have fine tuned many sense organs to their specific tasks with incredible perfection, sometimes reaching the limits of the physically possible regarding sensitivity and selectivity, but also impressive with regard to the efficient use of materials and energy. Biological sensory systems thus have an enormous potential for technical, industrial and medical applications. This applies to sensors specialized for many different forms of energy such as optical, electrical, magnetic, mechanical and chemical, for the detection of light, electrical and magnetic fields, sound, vibrations, motion, pressure, strain and stress, to name a few. Animals have many sensory capacities alien to our human experience, examples being sensitivities in the ultraviolet, infrared and ultrasound, electro-magnetic reception, and the detection of skeletal strain.

Simultaneously, remarkable advances in the areas of *synthetic materials and fabrication techniques* are making it increasingly possible to design and build highly sensitive, selective, and relatively inexpensive micro-/nano-electromechanical devices finding important sensory applications in a number of areas including: security (detection of air- and water-borne toxic materials); fluid motion sensing (in, for example, autonomous flying or underwater vehicles); medicine (implantable drug delivery systems, analysis of micro-/nano-liter chemical solutions, surgical tactile gloves); the industry (nano-liter amplification of DNA).

In order to make *effective use of all these innovation potentials*, and to develop new research strategies and new ways of thinking, it is of prime importance to intensify the dialogue among biologists, engineers, physical scientists, physicians and mathematicians, and to promote their actual collaboration and cross-disciplinary creativity. In this regard, the demands of industry, medicine, and government agencies with respect to the development and application of sensor arrays, data reduction, evaluation and com-

munication, and actuation decisions, all in the presence of background noise, are numerous and increasing in importance.

The *main goal of this book* is to promote the interaction between biologists, engineers, physical scientists, physicians and mathematicians at the frontiers of research and to help prepare the ground for innovative lines of highly multi-disciplinary future work. Our authors represent a wide spectrum of individuals from academia, government laboratories, and industry. As seen from the list of contents the topics treated cover a broad spectrum of problems ranging from processes of energy transformation and transduction to sensor array fabrication and application. These different fields are all linked and glued together by what a sensory system has to accomplish, both in biology and engineering. The main sections of the book are dedicated to sensory systems and capacities dealing with different forms of stimulus energy, such as vision, olfaction, various forms of mechanoreception, and infrared and electroreception. Importantly, there is an additional large section on bio-inspired synthetic sensors, sensor materials and the fabrication of technical sensors.

We acknowledge with pleasure and gratitude the funding in support of an ambitious international conference held in Cetraro/Calabria, Italy, in October 2008, whose success supported the idea of publishing a book on the same topic. Among the present authors many participated in this conference which was generously supported by the National Science Foundation, the University of Vienna (Faculty of Life Sciences), the US Air Force Office of Scientific Research, the University of Virginia at Charlottesville, the US Office of Naval Research, and the Engineering Conferences International. Thank you for sharing our vision. We also gratefully acknowledge the cooperation and patience of our authors and of Eva Maria Oberhauser of the editorial staff of Springer-Verlag, Vienna.

It is with great sadness that we had to accept the passing away of our dear friend Pepe Humphrey early in 2010, at a point in time when the concept of our book had already been worked out and the authors had already started to submit their contributions. The editorial work with the manuscripts, however, was still lying ahead of us. FGB did his best to take care of Pepe's share of the work. Pepe will be remembered by many as an outstanding person who not only very successfully promoted research collaborations between engineering and biology but also was a wonderful, warm hearted human being with an exceptionally broad mind, thinking deep and fully enjoying life.

Vienna, Charlottesville VA, Brisbane,
May 2011

Friedrich G. Barth
Joseph A. C. Humphrey†
Mandyam V. Srinivasan

Contents

III. Olfaction

IV. Mechanoreception

A. Hearing

B. Touch

C. Medium motion

D. Strain and substrate motion

V. Infrared and electro-reception

VI. Bioinspired sensors, sensor materials and fabrication

General

I

Nature as a model for technical sensors

1

Horst Bleckmann, Adrian Klein, Gunnar Meyer

Contents

Abstract

Nature has developed a stunning diversity of sensory systems. Humans rely mainly on visual information for object detection, discrimination and spatial orientation. In addition, they use acoustic, olfactory, and somatosensory cues. But even beyond these common sensory systems a large variety of highly specialized sensors have evolved in the animal kingdom. Examples include the infrared receptors of snakes and pyrophilous insects, the electroreceptors of fish, the magnetoreceptors of birds and the lateral line of fish and amphibians. This chapter deals with certain aspects of the detection and processing of hydrodynamic information in both natural and artificial lateral line systems. We show that the study of seemingly exotic sensory systems, such as the fish lateral line, can lead to discoveries that are useful for the construction of man-made sensors.

Horst Bleckmann
University of Bonn, Institute of Zoology
Poppelsdorfer Schloß, 53115 Bonn, Germany
e-mail: bleckmann@uni-bonn.de

1. Introduction

Animals use a stunning diversity of sensory systems to extract information from their environment. Most diurnal animals rely heavily on visual information in addition to using acoustic and chemosensory cues. For many nocturnal animals other sensory systems may be more important, such as mechanoreception (Bleckmann 1994), magnetoreception (Lohmann 2000), infrared reception (Ebert et al. 2007; Schmitz et al. 2007) or electroreception (von der Emde 1990; Manger and Pettigrew 1995).

For decades biologists have been studying the various types of sensory systems found in the animal kingdom. This helped gain insight in the uptake, transformation and transduction mechanisms these systems employ, in the behavior they mediate, and in the functional properties of their central pathways.

In general, biological sensors are miniaturized matched filters frequently of outstanding sensitivity (Wehner 1987). Parallel sampling and the processing of information is one of the ways nature improves the signal-to-noise ratio and reduces the likelihood of errors because of the malfunctioning of single sensory elements. Sensory receptors are connected to brain circuits that are specialized processors of information, with a speed unrivalled by our fastest computers. It is no wonder that the study of biological sensors, including their non-nervous auxiliary structures, has lead to important discoveries of interest to engineers and biologists alike. As a result, engineers have successfully emulated natural sensory systems in various types of technical applications, examples of which include retina-like sensors (Sandini and Metta 2003), electronic noses (Settles et al. 2003), silicon cochleae (Sarpeshkar 2003), spider trichobothria (Humphrey and Barth 2008), and slit sense organs (Barth 1978; see also chapter IV,5). This chapter focuses on the fish lateral line, a system consisting of arrays of mechanoreceptors, which has recently captured the curiosity of biologically inspired engineers (Fan et al. 2002; Peleshanko et al. 2007).

2. Sensory systems as matched filters

If we judge the performance of a sensory system we should keep in mind that this system was built by nature to fulfill a particular, restricted purpose. Although this restriction limits the amount of environmental information an animal can obtain, it eases the task of information processing. Sensory systems are shaped by natural selection, hence they tend to find only approximate solutions that are restricted to the narrow range of biologically relevant situations (for good examples, see Wehner 1987). To keep out unnecessary and/or to enhance relevant information, stimulus uptake begins with non-nervous auxiliary structures that allow only the energy of a particular form to be passed on. Examples include the visual, acoustic and chemosensory systems of invertebrates and vertebrates, the trichobothria and slit sense organs of spiders, the electrosensory systems of fishes and the infrared sense organs of pyrophilous insects and snakes. Though the information that reaches the receptor cells has already been filtered, sensory systems usually have several types of sensory cells that respond to different aspects of an adequate stimulus. For instance, the sensory cells in the tuberous organs of weakly electric fish may (B-cells) or may not (A-cells) be highly sensitive to stimulus phase (von der Emde and Bleckmann 1992). Since A- and B-cells are innervated by different afferent fibers that project to different brain areas, parallel processing in the electrosensory pathway already begins in the periphery. This is a common feature found in most, if not all, sensory systems.

3. Design rules of sensory systems

Although most animals are equipped with a variety of sensory systems, the cell types that are available for building these systems are fairly conservative. Instead of designing new systems from scratch, evolution redesigns already existing systems to serve a completely new function. The most common way to build a novel sensory system is to connect a certain sensory cell type with new non-nervous auxiliary structures. Ex-

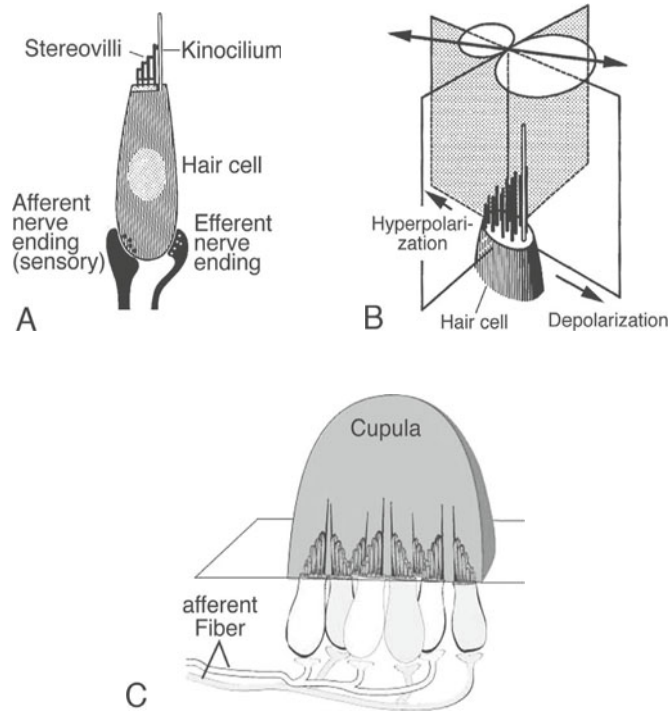


Fig. 1 A–C **A** A hair cell with a single kinocilium and several stereovilli (Flock 1971b). **B** The directional sensitivity of a hair cell. The membrane potential varies with the cosine of the angle between the direction of maximum sensitivity and the direction of the applied displacement stimulus (Flock 1971a). **C** A lateral line neuromast with antagonistically aligned hair cells, a cupula and two afferent nerve fibers. Efferent innervation is not shown

amples are the infrared receptors of the pyrophilous beetle *Melanophila* that originate from mechanosensory hairs (Schmitz et al. 2007). The inner ear of mammals and birds contains sensory systems that respond to sound pressure (the cochlea), to gravity (the otolith organs) and to angular acceleration (the semicircular canals). Despite the profound differences in function all three systems are based upon the mechanosensory hair cell. Hair cells have a hair bundle at their apical surface. This bundle is comprised of stereovilli that grow longer from one edge of the bundle to the other. A single true kinocilium (not present in all hair cells) occurs eccentrically at the tall edge of the hair bundle (Fig. 1A). This has functional consequences: deflecting the stereovilli towards the kinocilium causes a depolarization, while deflecing them away from the kinocilium results in a hyperpolarization of the cell membrane potential (Fig. 1B) (Flock and Wersäll 1962).

Fish also have otolithic organs that respond to gravity and semicircular canals that respond to angular acceleration. The fish ear lacks, however, a cochlea. For that reason, early researchers assumed that all fish were deaf. More recently though, careful studies have revealed that many fish (e. g. the Otophysi, clupeids and mormyrids) are highly sensitive to sound (for a recent review, see Sand and Bleckmann 2008). In these fish the swimbladder is linked to the ear and thus acts as a sound pressure receiver. Fish cannot only hear but can also determine the direction and the distance of a sound source. They do so by exploiting the particle movements – picked up directly by the otolithic apparatus of the inner ear – that are associated with sound pressure waves (Sand and Bleckmann 2008).

4. The fish lateral line

In fish a specialized hair cell system, called the lateral line, serves a particular function: it is used for both the detection of water movements and pressure gradients. In most fish the lateral line is composed of two types of receptors: those located on the surface of the skin, termed superficial neuromasts (SN), and those located in sub-epidermal fluid-filled canals, termed canal neuromasts (CN). Pore-like openings usually connect the interior of the lateral line canals with the outside medium (Fig. 2). The sensory epithelium of a lateral line neuromast contains about 20 to 100 hair cells (in a few cases up to 10 000 hair cells) (Bleckmann et al. 1989; Denton and Gray 1989; Münz 1989). The hair bundles of these cells protrude into a gelatinous cupula that connects the bundles either with the water surrounding the fish, or with canal fluid. The hair cells of a lateral line neuromast are oriented in two opposing directions (Fig. 1C) (Flock and Wersäll 1962). The information of SNs and CNs is relayed by separate nerve fibers to the brain (Münz 1985). A single nerve fiber contacts only hair cells with the same directional orientation, even if the fiber innervates more than one neuromast (Münz 1985).

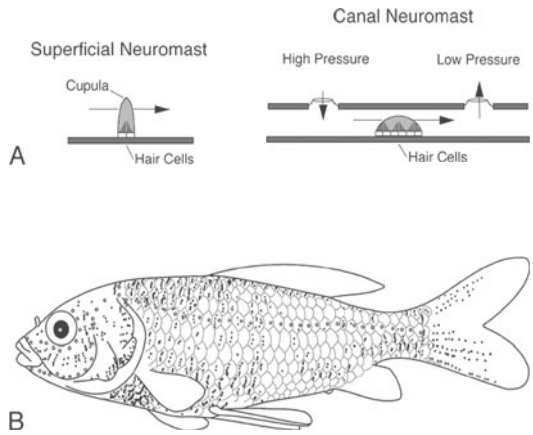


Fig. 2 **A**, **B** Morphology of the lateral line periphery. **A** Graphic representation of a SN and a lateral line canal with a CN (redrawn from Coombs and Montgomery 1999). Arrows indicate direction of water flow. **B** Distribution of SNs (dots) in the goldfish, *Carassius auratus*. Circles indicate canal pores. The drawing was gratefully provided by A. Grotefeld

There are about 30 000 extant fish species (Nelson 1984), which may live in very different hydrodynamic environments (e. g. in fast running rivers, along the ocean shoreline or in quiet lakes) and may have different ways of living (e. g. herbivorous, planctivorous or piscivorous). Therefore, it is not surprising that the organization of the peripheral lateral line can be quite diverse (e. g. Fig. 2B, 3). Differences include the number and distribution of SNs (Puzdrowski 1989; Schmitz et al. 2008), cupula dimensions (Trump and McHenry 2008) and the number, design and placement of lateral line canals (Denton and Gray 1983, 1988a; Bleckmann and Münz 1988; Coombs et al. 1988; Webb 1989; Weber et al. 1991). It is quite likely that the overall arrangement of the peripheral fish lateral line solves particular sensory problems, but so far a clear form-function relationship of the lateral line has not been uncovered for any fish species. However, several studies have already shown a strong link between the morphology of lateral line canals and their filter properties (Denton and Gray 1988b, 1989; van Netten and Khanna 1993; van Netten and Wiersinga-Post 2002; van Netten 2006).

4.1 Behavior

Fish use lateral line information for prey detection, predator avoidance, intraspecific communication, schooling, object detection and discrimination, entrainment and rheotaxis (Bleckmann 1993, 1994; Montgomery

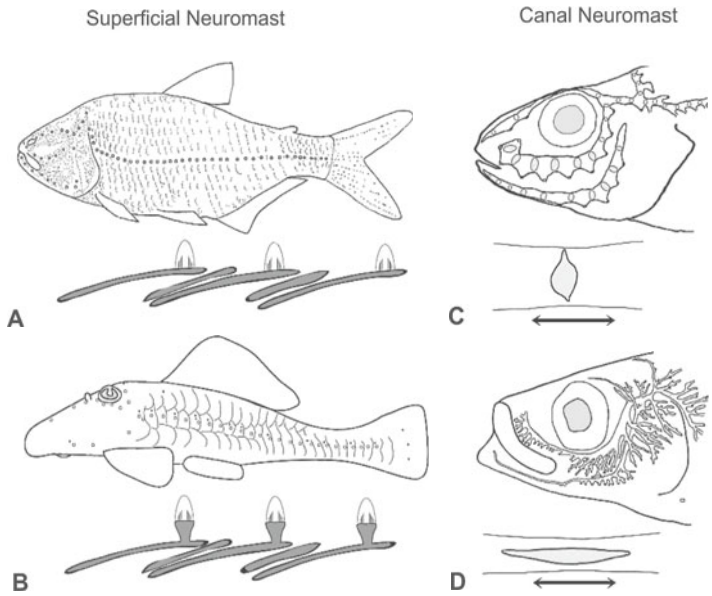


Fig. 3 **A–D** Examples of the morphological diversity of the fish lateral line. The peripheral lateral line of the blind cave fish, *Astyanax hubbsi* (redrawn from Schemmel 1967) **A** and the running water fish *Ancistrus* spec. (Grotefeld unpubl.) **B** Dots indicate SNs, circles represent canal pores. Note that *Astyanax* has much more SNs than *Ancistrus*. **C** The head of the freshwater fish *Percarina demidoffi*. Note the widened head lateral line canals (Jakubowski 1967). **D** Head of a herring *Sprattus sprattus*, a fish that lives in the open ocean. The black lines indicate the course of the lateral line canals (Blaxter et al. 1983)

et al. 1997; Liao 2007). However, it is primarily used for close range detection. For instance, blinded sculpins can estimate the distance from a dipole source from up to one body length away. Surface feeding fish respond to the capillary surface waves caused by prey insects trapped by the water surface from a distance of up to three body lengths away (Schwartz 1970; Hoin-Radkovski et al. 1984). For distance determination they use the dampening and dispersion properties of water surface waves (Bleckmann et al. 1989). The midwater fish *Carassius auratus* uses lateral line input to discriminate the direction of object motion, object speed, size and shape (Vogel and Bleckmann 2000). The blind cavefish *Astyanax mexicanus* (Characidae) readily passes through a barrier of rods without touching them. *Astyanax* makes use of the fact that any near-by stationary object interacts hydrodynamically with its body and thus alters the pattern of lateral line stimulation in a predictable way (Campenhausen et al. 1981). *Astyanax* even uses hydrodynamic information to develop an inner map of its aquatic environment (Teyke 1989; Burt de Perera 2004). Some piscivorous fish can sense and even track the hydrodynamic trail caused by their prey (Pohlmann et al. 2001, 2004; Pohlmann 2003).

Despite the presence of ambient currents in many aquatic environments, most behavioral experiments designed to uncover the sensory capabilities of the lateral line have been conducted under still water conditions. In a few studies, however, researchers have investigated the effect of background flow on behavioral thresholds. Successive elevations in background flow caused only a marked decrease in sensitivity in fish with a reduced canal system (Kanter and Coombs 2003; Bassett et al. 2006). While background noise may interfere with the ability of fish to detect prey, rheophilic fish may use lateral line input to ease their life in turbid rivers and streams. Trout (*Salvelinus fontinalis, Salmo trutta*) prefer to remain in specific locations, which they only leave to seize pieces of drifting debris. They even use hydrodynamic information to maintain their position and to

capture energy from vortices while moving through turbulent flow (Sutterlin and Waddy 1975; Liao et al. 2003; Beal et al. 2006).

4.2 Physiology of the fish lateral line

Primary lateral line afferents show ongoing activity. If stimulated with sinusoidal water motions, they respond (threshold water displacement is about 0.02 µm peak-to-peak at 100 Hz) with phase locking and – at larger stimulus amplitudes – with an increase in action potential rate (Bleckmann and Topp 1981; Kroese and Schellart 1992). In fish that have both a canal system and SNs, two types of primary afferents can be distinguished. Those that innervate SNs respond (up to a frequency of about 60 Hz) in proportion to water velocity, those that innervate CNs respond (up to a frequency of about 100 Hz) in proportion to water acceleration (Kroese and Schellart 1987; Kalmijn 1988). Due to the innervation pattern of lateral line neuromasts (see above and Fig. 1C) and the directional sensitivity of lateral line hair cells, primary lateral line afferents are directionally sensitive and phase lock only to one half of a full wave cycle.

Lateral line afferents not only respond to "clean" sinusoidal water motions, but also to the irregular water motions generated by the leg movements of planktonic prey (Montgomery and Macdonald 1987), to vortex rings that pass the fish laterally (Chagnaud et al. 2006), or to the Kármán vortex street caused by a rod exposed to running water (Chagnaud et al. 2007a). Lateral line afferents predictably respond to the well defined low frequency transient water motions, caused by an object that passes the fish laterally, with a single peak of excitation followed by a decrease in neural activity or vice versa. Due to the directional sensitivity of the lateral line hair cells, this response pattern inverses when the direction of object motion is reversed (Bleckmann and Zelick 1993; Mogdans and Bleckmann 1998; Mogdans and Geisen 2009). Afferents that are believed to innervate SNs additionally discharge numerous unpredictable bursts of action potentials in response to the irregular water motions that occur after the object has passed the fish (Mogdans and Bleckmann 1998).

As opposed to primary lateral line afferents, many central nervous lateral line units are highly selective. For instance, approximately 40 % of all medullary (medial octavolateralis nucleus) and midbrain (torus semicircularis; for central lateral line pathways see McCormick 1989 and Wullimann 1998) lateral line units do not respond to a small stationary sphere that vibrates with constant frequency and amplitude (Plachta et al. 1999; Bleckmann 2007). These units might, however, readily respond to amplitude modulated sinusoidal water motions (Plachta et al. 1999) or to the water motions caused by an object that passes the fish laterally. Units of this type may be highly directionally sensitive, i. e. they respond only if the object passes the fish in a certain direction (Bleckmann and Zelick 1993; Müller et al. 1996; Wojtenek et al. 1998). Other central lateral line units respond only to a stationary vibrating sphere, but not to a moving object (Engelmann and Bleckmann 2004).

4.3 Dipole localization

Sensory systems are usually built to detect, discriminate and localize a stimulus source. Localization of the source of a hydrodynamic stimulus is probably achieved by analyzing the responses of arrays of SNs and CNs. Both, physiological studies and mathematical calculations have shown that up to a distance of one body length the excitation patterns of arrays of lateral line neuromasts contain the information about the position

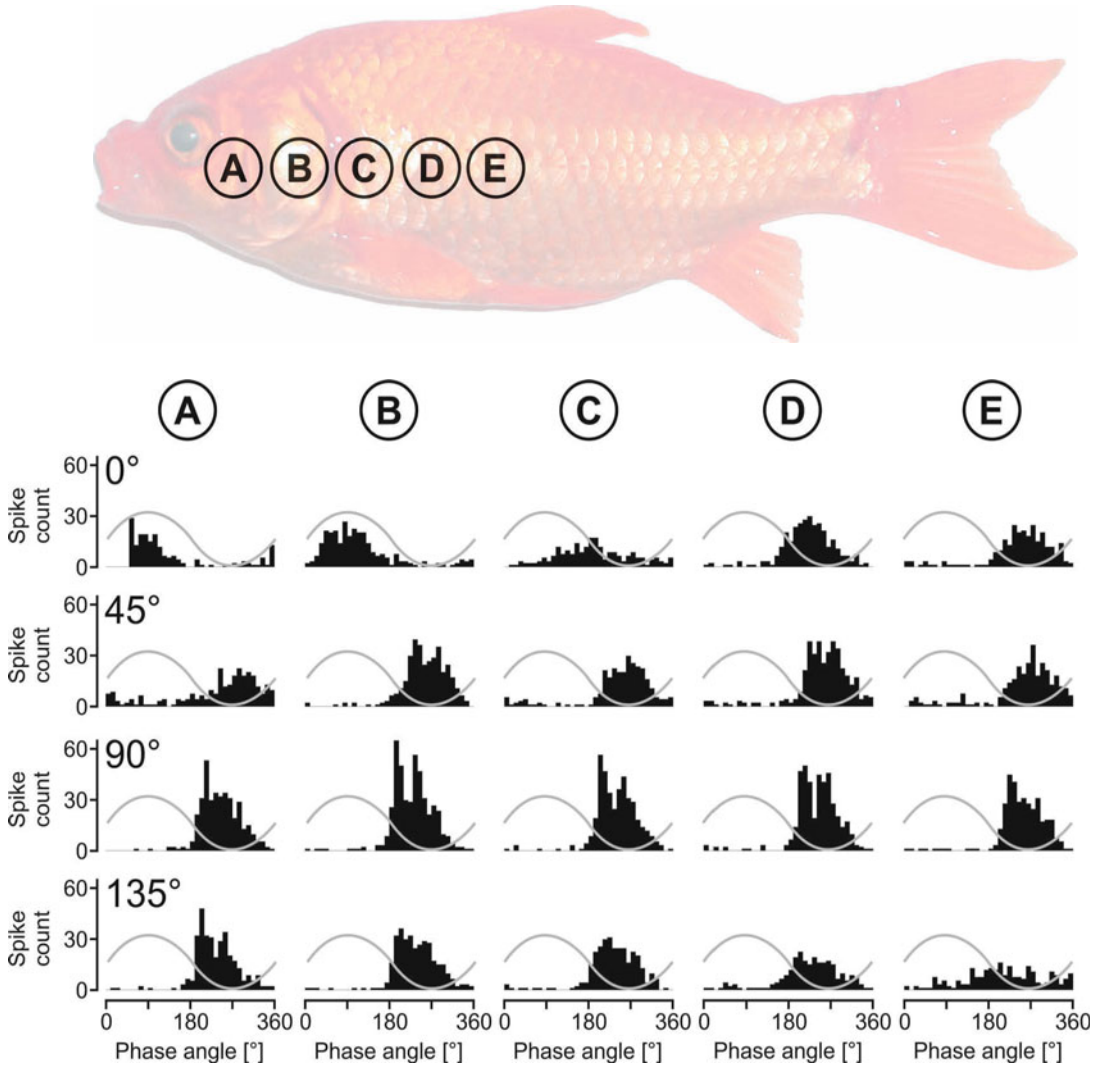


Fig. 4 Responses of a lateral line unit in the torus semicircularis of goldfish (top) to sinusoidal water motions generated by a vibrating sphere. Letters A to E denote sphere positions that were altered in 5 mm steps. Bottom: Period histograms (bin width 0.5 ms) of the neural activity induced by a stationary vibrating sphere (diameter 10 mm, vibration frequency 50 Hz, peak-to-peak displacement amplitude 220 µm, distance from fish 10 mm). Sphere vibration direction was 0° (parallel to the long axis of the fish), 45°, 90° (perpendicular to the long axis and the surface of the fish), and 135°. Note that the unit responded only to one half of a full wave cycle and that the phase angle of the response depended on both the position of the sphere and the direction of sphere vibration

(Sand 1981; Coombs et al. 2000) and the vibration direction (Sand 1981; Coombs et al. 2000; Curcic-Blake and van Netten 2006; Goulet et al. 2008) of the sphere. One of the uncovered stimulus parameters that is encoded in the central lateral line pathway is the direction of object motion; that is, some central lateral line units require a certain temporal pattern of water motions in order to respond (Müller et al. 1996; Wojtenek et al. 1998; Plachta et al. 2003). Although primary lateral line afferents transmit the information about object position and object vibration direction to the central nervous

system, space sensitive central units, like those known from the acoustic system of birds (Konishi 1986), have not been found in the brain of fish (Künzel, unpubl.; Meyer, unpubl.). Action potential discharge rates and/or phase angles of central lateral line neurons may, however, dramatically change if object position and/or object vibration direction is only slightly altered (e. g. Fig. 4). Thus, by monitoring the activity of many central lateral line units (population code), fish probably obtain the information necessary for object localization.

4.4 Running water

One often neglected but important question to ask is how hydrodynamic noise, e. g. bulk water flow, affects or is even used for lateral line perception. Bulk water flow consists of both a steady (DC) component and flow fluctuations, which are superimposed on the DC flow (Chagnaud et al. 2008 a). If exposed to unidirectional bulk water flow, flow sensitive lateral line afferents (most likely afferents that innervate SNs) tend to respond with a burst-like increase in ongoing activity (Voigt et al. 2000; Carton and Montgomery 2002; Engelmann et al. 2002; Chagnaud et al. 2007 b). Surprisingly, in nearly all flow-sensitive lateral line afferents this increase occurred irrespective of gross flow direction (Chagnaud *et al*. 2007 a). This suggests, that flow sensitive lateral line afferents do not respond to the DC component of the flow but only to the flow fluctuations (Chagnaud et al. 2007 b). Since flow fluctuations move with the mean flow, fish may determine gross flow direction and gross flow velocity by monitoring the direction and velocity of individual flow disturbances while these move across their body surface (Chagnaud et al. 2008 b). Some medullary (Kröther et al. 2002) and midbrain lateral line units (Zelick and Hofmann, unpubl.) have been found to be weakly tuned to velocity, but to date, lateral line units sharply tuned to water flow velocity have not been found in the medulla (Künzel 2009) or in the midbrain of fish (Bleckmann and Zelick 2009).

In running water (10 cm/s), the responses of flow sensitive primary lateral line afferents to the water motions caused by a stationary vibrating sphere, are masked, both in terms of discharge rate and phase locking. In contrast, responses of flow insensitive lateral line afferents evoked by a vibrating sphere are barely affected by running water (Engelmann et al. 2002; Chagnaud et al. 2007 b), provided the flow velocity is ≤ 10 cm/s (Chagnaud et al. 2007 b). If units of the medullary octavolateralis nucleus (MON) are stimulated with a stationary vibrating sphere in the presence of background flow, four types of units (MI to MIV) can be distinguished. MI units respond to running water with either an increase or a decrease in neural activity. If stimulated with a vibrating sphere, the response rates and/or the degree of phase-coupling decrease under running water conditions. MI units most likely receive excitatory or inhibitory (via inhibitory interneurons) input from afferents that innervate SNs. Ongoing discharge rates of MII units are hardly altered by running water. Moreover, these units do not change their responses to dipole stimuli in the presence of water flow. MII units most likely receive input from CNs. Like MII units, the ongoing discharge rates of MIII units are not altered in running water. However, their responses to dipole stimuli presented in background flow are masked (Kröther et al. 2004). Probably, excitatory input to MIII units, mediated via flow insensitive CN afferents, is inhibited by SN afferents. Like MI units, MIV units are flow-sensitive. However, the responses of these units to a vibrating sphere stimulus are not affected by running water. This suggests that MIV units also receive input from both SNs and CNs (Kröther et al. 2004). In goldfish, lateral

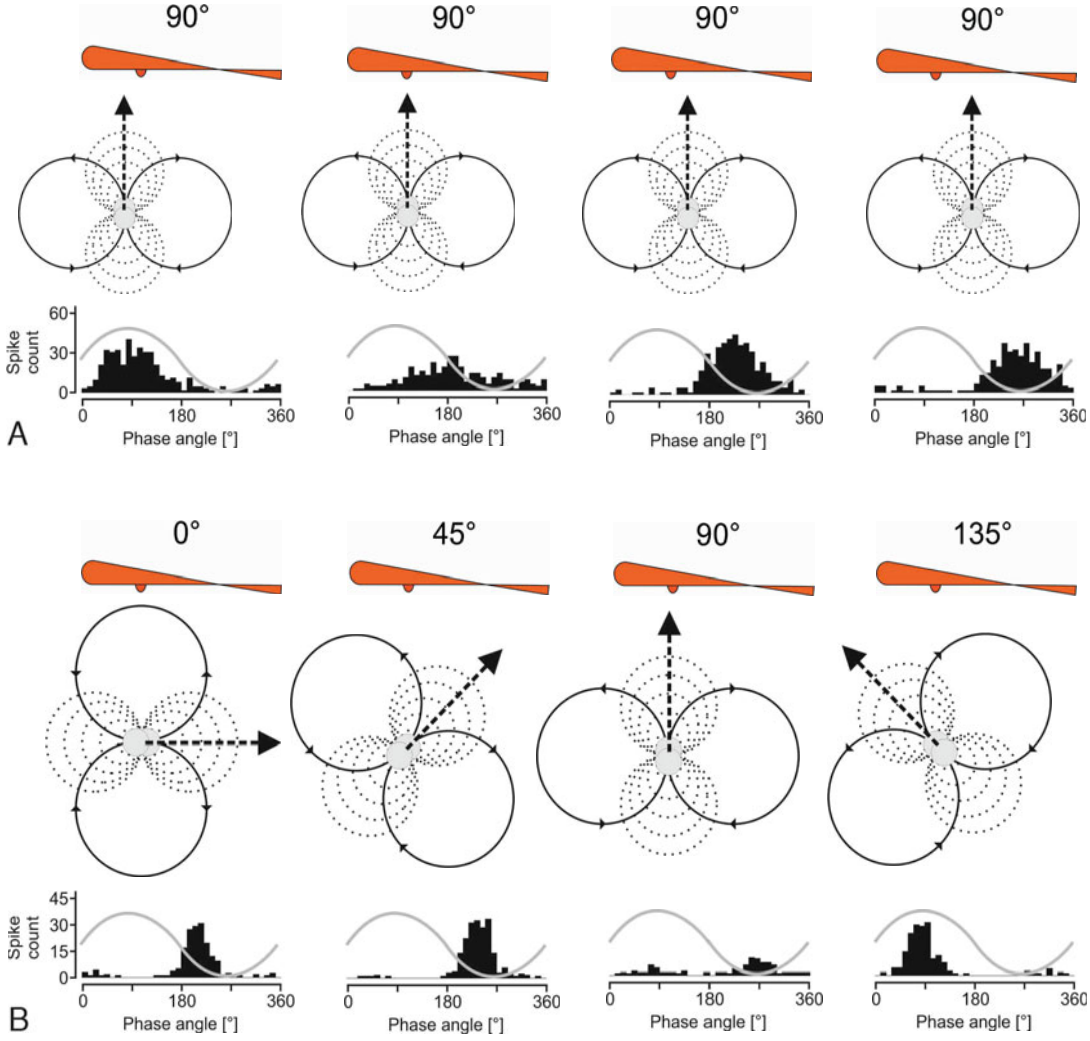


Fig. 5 **A, B** Iso-pressure contours (dashed lines) and flow lines (solid lines with arrows) around a dipole source. Iso-pressure contours are shown for the plane that bisects the source along its axis of oscillation, indicated by the large arrowheads. Note that for a given phase of sphere movement the flow direction on the surface of the fish depends on the direction of sphere vibration. **A** A change in the rostro-caudal position of the sphere by only 1 cm (sphere vibration direction 90°) caused a phase shift of about 180° in the neuronal response. **B** A change in the direction of sphere vibration from 45° to 135° also caused a 180° phase shift of the response (Meyer and Bleckmann, unpubl.). During recording the position of the sphere was not altered

line units recorded from the midbrain torus semicircularis are also either flow sensitive or insensitive (flow speed 10 cm/s), i. e. toral units receive either input from flow sensitive or from flow insensitive units of the MON. 90 % of all flow-insensitive toral lateral line units respond with a short burst if an object passes the fish, i. e. these units probably receive input from CNs. Like MON units, flow sensitive toral units show sustained increases or decreases in ongoing discharge rate if exposed to water flow. In still water, units

of both types may respond to a vibrating sphere and/or to a sphere passing the fish laterally (Engelmann and Bleckmann 2004). In running water, responses of flow-sensitive units to the vibrating sphere are masked (Engelmann and Bleckmann 2004). Interestingly, the responses of 12 out of 14 flow insensitive units were also masked by laminar water flow (Engelmann and Bleckmann 2004). Most likely, units of this type receive excitatory input from type MIII units. Consequently, as in the MON, a complete separation of SN and CN-input does not exist at the level of the torus semicircularis. In summary, MON and toral lateral line units display many different response types, especially under running water conditions. Unfortunately, the meaning of the various response types is still not fully understood. Moreover, the underlying neuronal circuits have up to now not been uncovered. It is hoped that intracellular recordings and neuroanatomical studies will help to bridge this gap.

The effects of running water on the responses of MON and toral lateral line units to a moving object are diverse and depend on object motion direction. This may be explained by peripheral hydrodynamic effects (Engelmann and Bleckmann 2004; Kröther et al. 2004). For a more detailed discussion on lateral line physiology the reader is referred to recent reviews (Bleckmann 2006, 2007; Bleckmann and Zelick 2009).

5. Biologically inspired hydrodynamic sensors

Numerous types of engineered flow sensors have been used for the measurement of surface and subsurface water motions (see also chapter VI,6 by Izadi and Krijnen). Examples include hot-wire anemometers, optoelectronic movement detectors, laser Doppler anemometers, capacitive probes, micro-pillars and custom-made underwater motion transducers (Bleckmann 1994; Chaves and Hodos 1998; Mogdans et al. 2004; Brücker et al. 2007; Brücker and Chaves 2007). Campenhausen et al. (1981) were the first to build an artificial lateral line neuromast. Their goal was to learn more about the hydrodynamic stimulus that reaches a blind cave fish when passing a nearby object. Denton and Gray (1983) and Weber et al. (1991) studied the functional properties of simple and complex artificial later line canals. These studies revealed that below about 100 Hz the water displacement in the canals is roughly proportional to the velocity of the outside medium (Denton and Gray 1983). Fan et al. (2002) were the first to build microfabricated artificial hair cell sensors. Arrays of artificial neuromasts can be used to detect and localize a dipole source (Yang et al. 2006). Lastly, an autonomous underwater vehicle equipped with arrays of artificial neuromasts can identify the signature of a wake and thus in principal can be trained to follow a wake generator (Yang et al. 2006).

All early researchers investigated the functional properties of artificial lateral line canals (ALLC) under still water conditions. By exposing an ALLC to running water, researchers learned that such a canal can be used to detect and localize a cylinder (object) placed adjacent to the canal (Pillapakkam et al. 2007). Such a cylinder induces complicated patterns of DC and AC fluid motions external to and inside the ALLC. These flow patterns depend on both the size and position of the cylinder with respect to the long axis of the ALLC. A narrow spread of flow oscillations inside the ALLC shows that there is little mechanical coupling between neighboring sections of the canal; that is, each short section of the canal behaves almost independently of its neighboring sections. Thus, both the

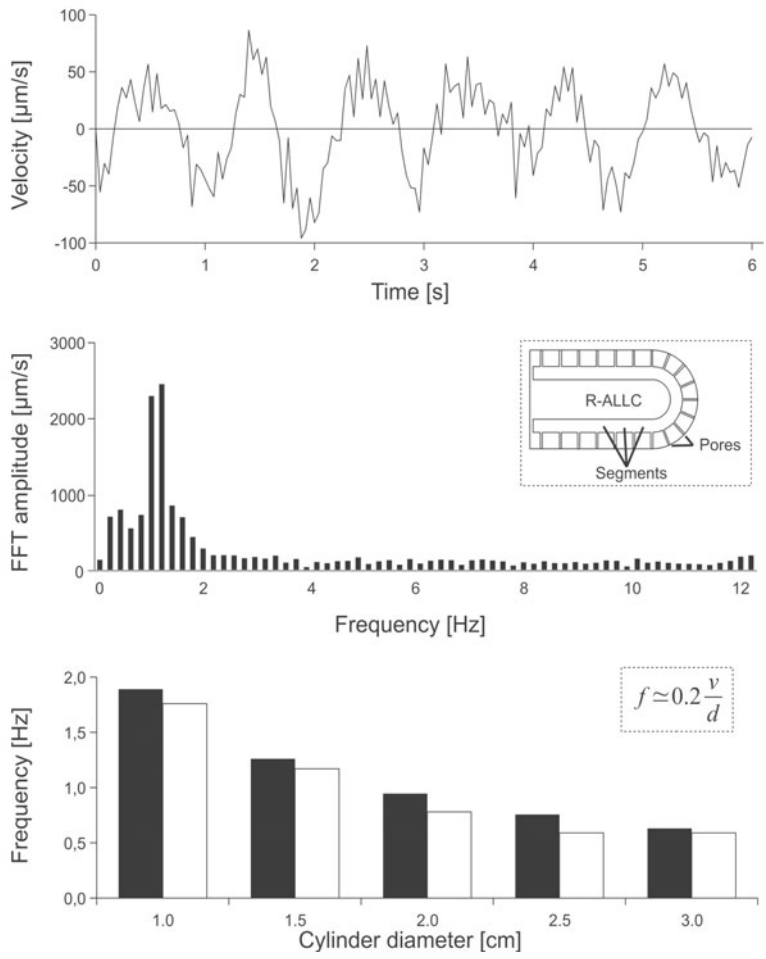


Fig. 6 A round artificial lateral line canal (R-ALLC) (inset in middle panel) was exposed to a cylinder placed upstream to the R-ALLC in the flow tank. The water inside the front segment of the R-ALLC oscillated with the vortex shedding frequency of the cylinder (top and FFT spectrum in the middle panel). Bottom: By using cylinders of different diameter vortex shedding frequencies can be altered. Peak oscillation frequencies of water motions inside the R-ALLC (white bars) were similar to the calculated (see equation in bottom panel) vortex shedding frequencies (black bars)

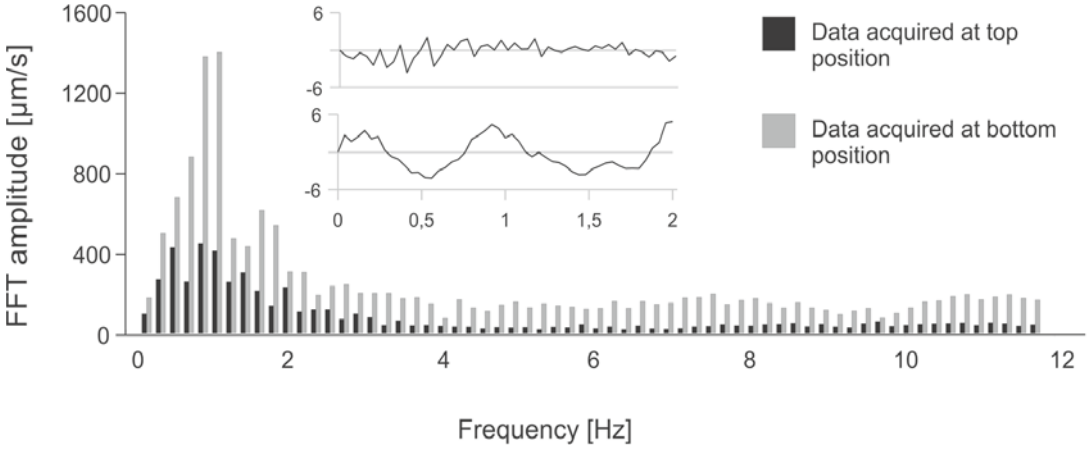


Fig. 7 FFT spectra of the flow velocities (inset) inside a vertically oriented R-ALLC. The spectrum, which is based on the measurements obtained from the upper part of the R-ALLC, is represented by black bars; the spectrum, which is based on the measurements obtained from the lower part of the R-ALLC, is represented by gray bars. In this experiment a horizontal cylinder was positioned upstream to the R-ALLC close to the bottom of the flow tank

flow and the flow oscillations in an ALLC not only indicate the presence, but also the relative position (with respect to the long axis of the canal) of an external object. Round ALLCs (R-ALLC, see Inset in middle panel of Fig. 6) can be used to determine the presence and position of a cylinder placed upstream to the ALLC. A cylinder that sheds vortices with a frequency of X Hz causes the water inside the R-ALLC to oscillate with X Hz (Fig. 6). While the frequency of these oscillations is identical with the vortex shedding frequency of the upstream cylinder, the amplitude of these oscillations depends on the position of the cylinder relative to the R-ALLC (c. f. Fig. 7). Because the vortex shedding frequency is not only a function of bulk flow velocity but also of cylinder diameter, the information about the size of a cylinder can also be obtained with an artificial lateral line, provided the bulk flow velocity is known.

Outlook

Hydrodynamic sensors that work according to the principles of the fish lateral line can potentially be used for a wide range of applications. Depending on the specific task, sensors can be designed for object detection, identification and localization as well as for flow measurements. Artificial lateral lines can also work under unfavorable conditions, such as darkness or in murky waters. Despite their astonishing capabilities, artificial lateral lines are still no match for natural lateral line systems. Therefore, the understanding of the biophysical and neural principles governing mechanosensation in fish remains an ongoing goal in scientific research (see also Chapter VI,6 by N Izadi and GJM Krijnen).

Acknowledgements

We thank Professor Friedrich Barth and Joseph Humphrey for the invitation to present a lecture at the meeting in Italy and to summarize parts of our work in this chapter. J. Mogdans and V. Schluessel deserve thanks for comments on an early version of the MS. Financial support came from DARPA, the DFG, the EU, and the BMBF.

References

Barth FG (1978) Slit sense organs: “Strain gauges” in the arachnid exoskeleton. Symposium Zool Soc London 42: 439–448

Bassett DK, Carton AG, Montgomery JC (2006) Flowing water decreases hydrodynamic signal detection in a fish with an epidermal lateral-line system. Marine Freshwater Res 57: 611–617

Beal DN, Hover FS, Triantafyllou MS, Liao JC, Lauder GV (2006) Passive propulsion in vortex wakes. J Fluid Mech 549: 385–402

Blaxter JHS, Gray JAB, Best ABC (1983) Structure and development of the free neuromasts and the lateral line system of the herring. J Mar Biol Ass UK 63: 247–260

Bleckmann H (1993) Role of the lateral line and fish behavior. In: Pitcher TJ (ed) Behaviour of teleost fishes. Chapman and Hall, London New York Tokyo, pp 201–246

Bleckmann H (1994) Reception of hydrodynamic stimuli in aquatic and semiaquatic animals. In: Rathmayer W (ed) Progress in zoology. Vol 41. Gustav Fischer, Stuttgart Jena New York, pp 1–115

Bleckmann H (2006) The lateral line system of fish. In: Hara T, Zielinski B (eds) Sensory systems neuroscience: Fish physiology, Vol 25. Elsevier, Amsterdam, pp 411–453

Bleckmann H (2007) Peripheral and central processing of lateral line information. J Comp Physiol A 194: 145–158

Bleckmann H, Münz H (1988) The anatomy and physiology of lateral line mechanoreceptors in teleosts with multiple lateral lines. In: Barth FG (ed) Verh Dtsch Zool Ges 81. Gustav Fischer, Stuttgart, pp 288

Bleckmann H, Tittel G, Blübaum-Gronau E (1989) The lateral line system of surface-feeding fish: Anatomy, physiology, and behavior. In: Coombs S, Görner P, Münz H (eds) The mechanosensory lateral line. Neurobiology and evolution. Springer, New York, pp 501–526

Bleckmann H, Topp G (1981) Surface wave sensitivity of the lateral line organs of the topminnow *Aplocheilus lineatus*. Naturwissenschaften 68: 624–625

Bleckmann H, Zelick R (1993) The responses of peri-

pheral and central mechanosensory lateral line units of weakly electric fish to moving objects. J Comp Physiol A 172: 115–128

Bleckmann H, Zelick R (2009) Lateral line system of fish. Integrative Zool 4: 13–25

Brücker C, Bauer D, Chaves H (2007) Dynamic response of micro-pillar sensors measuring fluctuating wall-shear-stress. Exp Fluids 42: 737–749

Burt de Perera T (2004) Spatial parameters encoded in the spatial map of the blind Mexican cave fish, *Astyanax fasciatus*. Animal Behav 68: 291–295

Campenhausen Cv, Riess I, Weissert R (1981) Detection of stationary objects in the blind cave fish *Anoptichthys jordani* (Characidae). J Comp Physiol A 143: 369–374

Carton AG, Montgomery JC (2002) Responses of lateral line receptors to water flow in the Antarctic notothenoid, *Trematomus bernacchii*. Polar Biol 25: 789–793

Chagnaud BP, Bleckmann H, Engelmann J (2006) Neural responses of goldfish lateral line afferents to vortex motions. J Exp Biol 209: 327–342

Chagnaud BP, Bleckmann H, Hofmann M (2007a) Kármán vortex street detection by the lateral line. J Comp Physiol A 193: 753–763

Chagnaud BP, Bleckmann H, Hofmann MH (2008a) Lateral line nerve fibers do not respond to bulk water flow direction. J Zool 111: 204–217

Chagnaud BP, Brücker C, Hofmann MH, Bleckmann H (2008b) Measuring flow velocity and flow direction by spatial and temporal analysis of flow fluctuations. J Neurosci 28: 4479–4487

Chagnaud BP, Hofmann MH, Mogdans J (2007b) Responses to dipole stimuli of anterior lateral line nerve fibres in goldfish, *Carassius auratus*, under still and running water conditions. J Comp Physiol A 193: 249–263

Chaves LM, Hodos W (1998) Color reversal-learning deficits after tectofugal pathway lesions in the pigeon telencephalon. Behav Brain Res 90: 1–12

Coombs S, Finneran JJ, Conley RA (2000) Hydrodynamic imaging formation by the lateral line system of the Lake Michigan mottled sculpin, *Cottus bairdi*. Phil Trans R Soc B 355: 1111–1114

Coombs S, Janssen J, Webb JF (1988) Diversity of lateral line systems: evolutionary and functional considerations. In: Atema J, Fay RR, Popper AN, Tavolga WN (eds) Sensory biology of aquatic animals. Springer, New York, pp 553–593

Coombs S, Montgomery JC (1999) The enigmatic lateral line. In: Fay RR, Popper AN (eds) Comparative hearing: fish and amphibians. Springer, New York, pp 319–362

Curcic-Blake B, van Netten SM (2006) Source localization encoding in the fish lateral line. J Exp Biol 209: 1548–1559

Denton EJ, Gray JAB (1983) Mechanical factors in the excitation of clupeid lateral lines. Proc R Soc Lond B 218: 1–26

Denton EJ, Gray JAB (1988a) Mechanical factors in the excitation of lateral line canals. In: Atema J, Fay RR, Popper AN, Tavolga WN (eds) Sensory biology of aquatic animals. Springer, New York, pp 595–617

Denton EJ, Gray JAB (1988b) Mechanical factors in the excitation of the lateral lines of fishes. In: Atema J, Fay RR, Popper AN, Tavolga WN (eds) Sensory biology of aquatic animals. Springer, New York, pp 595–617

Denton EJ, Gray JAB (1989) Some observations on the forces acting on neuromasts in fish lateral line canals. In: Coombs S, Görner P, Münz H (eds) The mechanosensory lateral line. Neurobiology and evolution. Springer, New York, pp 229–246

Ebert J, Müller S, Westhoff G (2007) Behavioral examination of the infrared sensitivity of ball pythons. J Zool 272: 340–347

Engelmann J, Bleckmann H (2004) Coding of lateral line stimuli in the goldfish midbrain in still- and running water. J Zool 107: 135–151

Engelmann J, Hanke W, Bleckmann H (2002) Lateral line reception in still- and running water. J Comp Physiol A 188: 513–526

Fan Z, Chen J, Zou J, Bullen D, Liu C, Delcomyn F (2002) Design and fabrication of artifical lateral line flow sensors. J Micromech Microeng 12: 655–661

Flock A (1971a) Sensory transduction in hair cells. I. Principles of receptor physiology. In: Loewenstein WR (ed) Handbook of sensory physiology. Springer, New York, pp 396–441

Flock A (1971b) The lateral line organ mechanoreceptors. In: Hoar WS, Randall DJ (eds) Fish physiology, Vol. 5. Academic Press, New York, pp 241–263

Flock A, Wersäll J (1962) A study of the orientation of sensory hairs of the receptor cells in the lateral line organ of a fish with special reference to the function of the receptors. J Cell Biol 15: 19–27

Goulet J, Engelmann J, Chagnaud BP, Franosch J-MP, Suttner MD, van Hemmen JL (2008) Ob-

ject localization through the lateral line system of fish: theory and experiment. J Comp Physiol A 194: 1–17

Hoin-Radkovski I, Bleckmann H, Schwartz E (1984) Determination of source distance in the surface-feeding fish *Pantodon buchholzi* (Pantodontidae). Animal Behav 32: 840–851

Humphrey JAC, Barth FG (2008) Medium flow-sensing hairs: Biomechanics and models. In: Casas J, Simpson J (eds) Advances in insect physiology, Vol 34. Academic Press, London, pp 1–80

Jakubowski M (1967) Cutaneous sense organs of fishes. VIII. The structure of the system of lateral-line canal organs in the Percidae. Acta Biol Cracov Ser Zool 10: 69–81

Kalmijn AJ (1988) Hydrodynamic and acoustic field detection. In: Atema J, Fay RR, Popper AN, Tavolga WN (eds) Sensory biology of aquatic animals. Springer, New York, pp 83–130

Kanter MJ, Coombs S (2003) Rheotaxis and prey detection in uniform currents by lake michigan mottled sculpin (*Cottus bairdii*). J Exp Biol 206: 59–70

Konishi M (1986) Centrally synthesized maps of sensory space. Trends Neurosci 100: 163–168

Kroese ABA, Schellart NAM (1987) Evidence for velocity- and acceleration-sensitive units in the trunk lateral line of the trout. J Physiol 393: 29

Kroese ABA, Schellart NAM (1992) Velocity- and acceleration-sensitive units in the trunk lateral line of the trout. J Neurophysiol 68: 2212–2221

Kröther S, Bleckmann H, Mogdans J (2004) Effects of running water on brainstem latetral line responses in trout, *Oncorhynchus mykiss*, to sinusoidal wave stimuli. J Comp Physiol A 190: 437–448

Kröther S, Mogdans J, Bleckmann H (2002) Brainstem lateral line responses to sinusoidal wave stimuli in still- and running water. J Exp Biol 205: 1471–1484

Künzel S (2009) Characterisation of brainstem lateral line neurons in goldfish, *Carassius auratus*: Frequency selectivity, spatial excitation patterns and flow sensitivity. PhD thesis, University of Bonn, Germany.

Liao JC (2007) A review of fish swimming mechanics and behaviour in altered flows. Phil Trans R Soc Lond B 362: 1973–1993

Liao JC, Beal DN, Lauder GV, Triantafyllou MS (2003) The Kármán gait: Novel kinematics of rainbow trout swimming in a vortex street. J Exp Biol 206: 1059–1073

Lohmann KJ (2000) The neurobiology of magnetoreception in vertebrate animals. T Neurosci 23: 153–159

Manger PR, Pettigrew JD (1995) Electroreception and the feeding behaviour of platypus (*Ornithorhynchus anatius*: Monotremata: Mammalia). Phil Trans R Soc London 347: 359–381

McCormick CA (1989) Central lateral line mechanosensory pathways in bony fish. In: Coombs S, Görner P, Münz H (eds) The mechanosensory lateral line. Neurobiology and evolution. Springer, New York, pp 341–364

Mogdans J, Bleckmann H (1998) Responses of the goldfish trunk lateral line to moving object. J Comp Physiol A 182: 659–676

Mogdans J, Geisen S (2009) Responses of the goldfish head lateral line to moving objects. J Comp Physiol A 195: 151–165

Mogdans J, Kröther S, Engelmann J (2004) Neurobiology of the fish lateral line: Adaptations for the detection of hydrodynamic stimuli in running water. In: von der Emde G, Mogdans J, Kapoor GB (eds) The senses of fish. Adaptations for the reception of natural stimuli. Narosa Publishing House, New Delhi, pp 265–287

Montgomery JC, Baker CF, Carton AG (1997) The lateral line can mediate rheotaxis in fish. Nature 389: 960–963

Montgomery JC, Macdonald JA (1987) Sensory tuning of lateral line receptors in Antarctic fish to the movements of planctonic prey. Science 235: 195–196

Müller HM, Fleck A, Bleckmann H (1996) The responses of central octavolateralis cells to moving sources. J Comp Physiol A 179: 455–471

Münz H (1985) Single unit activity in the peripheral lateral line system of the cichlid fish *Sarotherodon niloticus* L. J Comp Physiol A 157: 555–568

Münz H (1989) Functional organization of the lateral line periphery. In: Coombs S, Görner P, Münz H (eds) The mechanosensory lateral line. Neurobiology and evolution. Springer, New York, pp 285–298

Nelson JS (1984) Fishes of the world. John Wiley and Sons, New York

Peleshanko S, Julian MD, Ornatska M, McConney ME, LeMieux MC, Chen N, Tucker C, Yang Y, Liu C, Humphrey JAC, Tsukruk VV (2007) Hydrogel-Encapsulated microfabricated hair cells mimicking fish cupula neuromast. Adv Mater 19: 2903–2909

Pillapakkam SB, Barbier C, Humphrey AC, Rüter

A, Otto B, Bleckmann H, Hanke W (2007) Experimental and numerical investigation of a fish artificial lateral line canal. In: 5th International Symposium on turbulence and shear flow phenomena. TU München, pp 1–6

Plachta D, Hanke W, Bleckmann H (2003) A hydrodynamic topographic map and two hydrodynamic subsystems in a vertebrate brain. J Exp Biol 206: 3479–3486

Plachta D, Mogdans J, Bleckmann H (1999) Responses of midbrain lateral line units of the goldfish, *Carassius auratus*, to constant-amplitude and amplitude modulated water wave stimuli. J Comp Physiol A 185: 405–417

Pohlmann K (2003) When the night comes: Non-visual predator-prey interactions in fish. Dissertation. University of Konstanz, Konstanz

Pohlmann K, Atema J, Breithaupt T (2004) The importance of the lateral line in nocturnal predation of piscivorous catfish. J Exp Biol 207: 2971–2978

Pohlmann K, Grasso FW, Breithaupt T (2001) Tracking wakes: The nocturnal predatory strategy of piscivorous catfish. Proc Nat Acad Sci 98: 7371–7374

Puzdrowski RL (1989) Peripheral distribution and central projections of the lateral-line nerves in goldfish, *Carassius auratus*. Brain Behav Evol 34: 110–131

Sand O (1981) The lateral line and sound reception. In: Tavolga WN, Popper AN, Fay RR (eds) Hearing and sound communication in fishes. Springer, New York, pp 459–480

Sand O, Bleckmann H (2008) Orientation to auditory and lateral line stimuli. In: Webb JF, Fay RR, Popper AN (eds) Fish bioacoustics, vol 22. Springer, New York, pp 183–232

Sandini G, Metta G (2003) Retina-like sensors: motivations, technology and applications. In: Barth FG, Humphrey JAC, Secomb TW (eds) Sensors and sensing in biology and engineering. Springer Verlag, Wien New York, pp 251–262

Sarpeshkar R (2003) The silocon cochlea. In: Barth FG, Humphrey JAC, Secomb TW (eds) Sensors and sensing in biology and engineering. Springer, Wien New York.

Schemmel C (1967) Vergleichende Untersuchungen an den Hautsinnesorganen ober- und unterirdisch lebender *Astyanax*-Formen. Z Morph Tiere 61: 255–316

Schmitz A, Bleckmann H, Mogdans J (2008) Organization of the superficial neuromast system in goldfish, *Carassius auratus*. J Morphol 269: 751–761

Schmitz A, Sehrbrock A, Schmitz H (2007) The analysis of the mechanosensory origin of the infrared sensilla in *Melanophila acuminata* (Coleoptera; Bupestridae) adduces new insight into the transduction mechanism. Arth Struct Develop 36: 291–303

Schwartz E (1970) Ferntastsinnesorgane von Oberflächenfischen. Z Morphol Tiere 67: 40–57

Settles GS, Kester DA, Dodson-Dreibelbis LJ (2003) The external aerodynamics of canine olfaction. In: Barth FG, Humphrey JAC, Secomb TW (eds) Sensors and sensing in biology and engeneering. Springer, Wien New York, pp 323–335

Sutterlin AM, Waddy S (1975) Possible role of the posterior lateral line in obstacle entrainment by brook trout (*Salvelinus fontinalis*). J Fish Res Bd. Canada 32: 2441–2446

Teyke T (1989) Learning and remembering the environment in the blind cave fish *Anoptichthys jordani*. J Comp Physiol A 164: 655–662

Trump WJV, McHenry MJ (2008) The morphology and mechanical sensitivity of lateral line receptors in zebrafish larvae (*Danio rerio*). J Exp Biol 211: 2105–2115

van Netten SM (2006) Hydrodynamic detection by cupulae in a lateral line canal: functional relations between physics and physiology. Biol Cybern 94: 67–85

van Netten SM, Khanna SM (1993) Mechanical demodulation of hydrodynamic stimuli performed by the lateral line organ. In: Allum JHJ, Allum-Mecklenburg DJ, Harris FP, Probst R (eds) Prog Brain Res. Elsevier, Amsterdam, pp 45–51

van Netten SM, Wiersinga-Post C (2002) Matched peripheral filtering in the lateral line organ and relation to temperature. Bioacoustics 12: 153–156

Vogel D, Bleckmann H (2000) Behavioral discrimination of water motions caused by moving objects. J Comp Physiol A 186: 1107–1117

Voigt R, Carton AG, Montgomery JC (2000) Responses of anterior lateral line afferent neurones to water flow. J Exp Biol 203: 2495–2502

von der Emde G (1990) Discrimination of objects through electrolocation in the weakly electric fish, *Gnathonemus petersii*. J Comp Physiol A 167: 413–421

von der Emde G, Bleckmann H (1992) Differential responses of two types of electroreceptive af-

ferents to signal distortions may permit capacitance measurement in a weakly electric fish, *Gnathonemus petersii*. J Comp Physiol A 171: 683–694

Webb JF (1989) Gross morphology and evolution of the mechanoreceptive lateral-line system in teleost fishes. Brain Behav Evol 33: 34–53

Weber T, Bleckmann H, Münz H (1991) Model experiments regarding the function of complex lateral line canals. In: Pfannenstiel H-D (ed) Verh Dtsch Zool Ges 84. Gustav Fischer, Stuttgart, pp 461–462

Wehner R (1987) Matched filters – neural models of the external world. J Comp Physiol A 161: 511–531

Wojtenek W, Mogdans J, Bleckmann H (1998) The responses of midbrain lateral line units of the goldfish *Carassius auratus* to moving objects. J Zool 101: 69–82

Wullimann MF (1998) The central nervous system. In: Evans DH (ed) The physiology of fishes. CRC Press, New York, pp 245–282

Yang Y, Chen J, Enge J, Pandya S, Chen N, Tucker C, Coombs S, Jones DL, Liu C (2006) Distant touch hydrodynamic imaging with an artificial lateral line. P Nat Acad Sci 103: 18891–18895

From biology to engineering: insect vision and applications to robotics

2

Mandyam V. Srinivasan, Richard J. D. Moore,
Saul Thurrowgood, Dean Soccol, Daniel Bland

Contents

Abstract

The past two decades have witnessed a growing interest not only in understanding sensory biology, but also in applying the principles gleaned from these studies to the design of new, biologically inspired sensors for a variety engineering applications. This chapter provides a brief account of this interdisciplinary endeavour in the field of insect vision and flight guidance. Despite their diminutive eyes and brains, flying insects display superb agility and remarkable navigational competence. This review describes our current understanding of how insects use vision to stabilize flight, avoid collisions with objects, regulate flight speed, navigate to a distant food source, and orchestrate smooth landings. It also illustrates how some of these insights from biology are being used to develop novel algorithms for the guidance of terrestrial and airborne vehicles. We use this opportunity to also highlight some of the outstanding questions in this particular area of sensing and control.

Mandyam V. Srinivasan
University of Queensland
Queensland Brain Institute
St. Lucia, QLD 4072, Australia
e-mail: m.srinivasan@uq.edu.au

1. Introduction

Flying insects, with their compound eyes, miniature brains, and amazingly aerobatic airframes, are capturing the interest of scientists and engineers from a variety of disciplines. The reasons for this interest are manifold. First, there is a desire to better understand the sensors, the processing algorithms and the strategies for flight control and navigation that enable a fly to evade a descending hand, or a honeybee to find a food source 10 km away from its hive and return home unerringly. Second, engineers are keen to draw inspiration from biology to design new generations of sensors, signal processors and flight controllers. Third, flying insects are living proof that it should eventually be possible to engineer airborne vehicles that are agile, navigationally competent, weigh less than 100 milligrams, consume about 60 milliwatts of power during flight, and can circumnavigate the world using the energy equivalent of an ounce of honey.

This chapter describes briefly the sensors and the principles of information processing that underlie visually guided flight in insects, and the application of some of these principles to achieve autonomous flight in aircraft. We will focus primarily on honeybees (*Apis mellifera*), as these are the creatures that have been investigated most extensively in this context.

1.1 The compound eye

The compound eyes of insects are obviously very different in appearance from the so-called "simple" eyes of vertebrates. While a vertebrate eye is composed of a single lens that focuses light from the outside world on to a sheet of photoreceptive neurons in the retina – rather like a camera – the compound eye of an insect comprises a large number of "little eyes", so-called "ommatidia", arranged on an approximately spherical surface (e. g. Goodman 2003, Borst 2009). Each ommatidium comprises a small lens that captures light coming in from a small patch of the environment, typically a few degrees in diameter, and focuses it onto a small group of 8 or 9 photoreceptors. In a honeybee, the two compound eyes together carry an array of 11,000 ommatidia *in toto*, enabling the insect to command a nearly all-round view of the environment in which each ommatidium collects information about the intensity, colour (and possibly other properties) of light arriving from a small patch of the environment that is about 2 deg in diameter (Srinivasan 2011). Thus, one might say that the compound eyes capture and represent the world as a panoramic, pointillistic image, with an optical system that is difficult to match, in terms of size and weight, using a vertebrate-style eye design.

1.2 The ocelli

In addition to the compound eyes, many insects possess three small, rudimentary, "simple" eyes, known as ocelli, located at the top of the head. One, termed the median ocellus, is positioned dorsofrontally and views the horizon in front of the flying insect. Two others, known as the lateral ocelli, are situated dorsolaterally and view the horizon to either side. A counter-clockwise roll of the head would cause the right lateral ocellus to see more of the sky and less of the ground, and therefore receive more light than the left lateral ocellus, which would experience the opposite effect. The difference between the optical stimulations received by the left and right ocelli can be used to sense and correct deviations in roll attitude. Similarly, the median ocellus receives more

or less light according to whether the insect pitches up or down. Therefore, the signals from the three ocelli, when processed appropriately, can provide robust information on the orientation of the insect's head relative to the horizon, and thus an accurate estimate of roll and pitch. Indeed, behavioural studies with tethered, flying dragonflies have shown that compensatory rolling and pitching movements of the head can be elicited by appropriate optical stimulation of the three ocelli, using light guides (Stange 1981).

1.3 Limitations of compound eyes

While the compound eyes of insects confer distinct advantages, they also impose challenges. Unlike vertebrates, insects have immobile eyes with fixed-focus optics. Therefore, they cannot infer the distance of an object from the extent to which the directions of gaze must converge to view the object, or by monitoring the refractive power that is required to bring the image of the object into focus on the retina. Furthermore, compared with human eyes, the eyes of insects are positioned much closer together, and possess inferior spatial acuity. Therefore, even if an insect possessed the neural apparatus required for binocular stereopsis, such a mechanism would be relatively imprecise and restricted to measuring ranges of only a few centimetres (Srinivasan 1993). Not surprisingly, insects have evolved alternative visual strategies for guiding locomotion and for "seeing" the world in three dimensions. Many of these strategies rely on using cues derived from the image motion that the animal experiences in its eyes when it moves in its environment, as we shall describe below.

2. Visual guidance of flight

2.1 Stabilization of flight direction and attitude

When an insect flying along a straight line is blown to the left by a gust of wind, the image on its frontal retina moves to the right. This causes the flight motor system to generate a corrective turning response – the so-called "optomotor response" – that brings the insect back on course (Reichardt 1969). By evaluating the direction of motion of the panoramic image of the world in the eye, the flying insect is able to stabilize unwanted rotations not only in yaw, but also in roll and pitch. The lobula plate, which represents the fourth stage of processing in the visual pathway of the fly, carries a number of large-field, motion sensitive neurons that are selectively responsive to rotations of the head about the yaw, pitch and roll axes, as well as other axes (Borst 2009, Krapp 2000, Egelhaaf 2008, Joesch et al.2008, Nordstrom et al. 2008).Thus, the ensemble of responses across these neurons provides accurate information about the direction of head rotation and drives downstream motorneurons, involved in flight control, to generate appropriate corrective rotations of the head and body.

2.2 Avoiding obstacles and negotiating narrow gaps

When an insect flies in a straight line, the image of a given point in the scene will move in the eye at a speed (angular speed, measured in degrees per second) that is proportional to the speed of flight and inversely proportional to the distance to the point (Fig. 1a). Thus, distant objects in the scene will induce low image speeds, whereas nearby objects will generate rapid image motion.

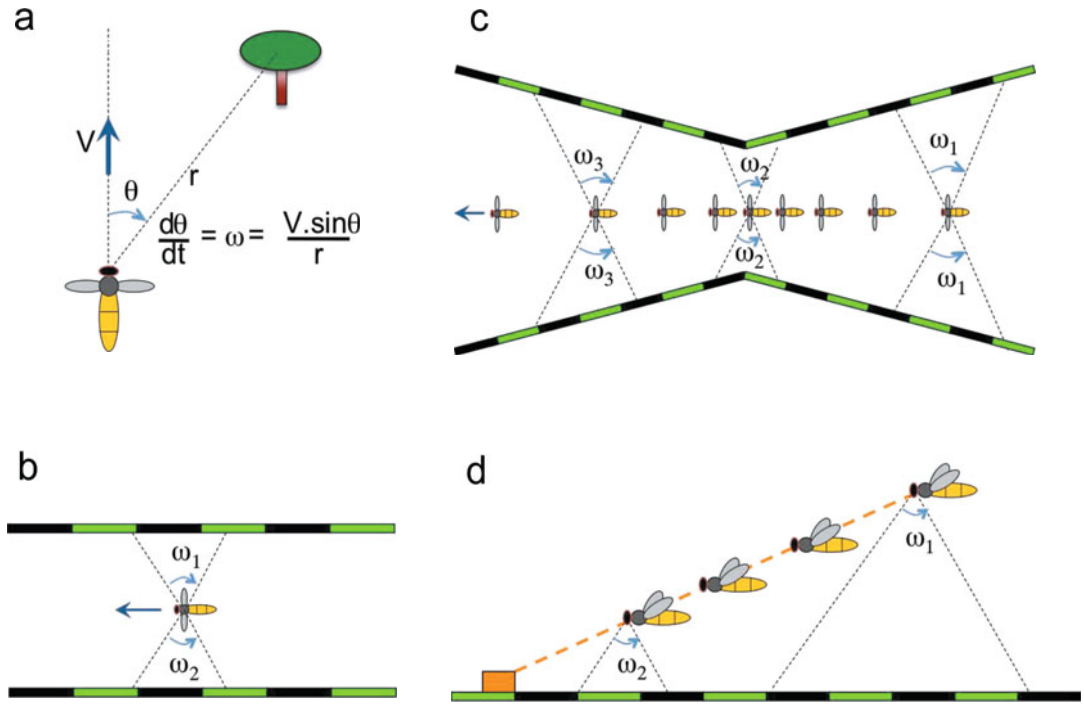


Fig. 1 **a** Illustration of how the angular velocity (omega) of the image of an object in the eye depends upon an insect's flight speed (V), as well as the range (r) and bearing (θ) of the object. **b** Narrow passages are negotiated safely by flying a trajectory such that both eyes experience the same image velocity in the lateral fields of view (ω_1=ω_2). **c** Flight speed is controlled by holding the image velocities in the lateral fields of view constant throughout the flight ($\omega_1 = \omega_2 = \omega$). **d** Smooth landings are executed by decreasing flight speed progressively such that the image velocity in the ventral field of view is held constant ($\omega_1 = \omega_2$). Adapted with permission from (Srinivasan, in press)

Experiments using moving stimuli reveal that honeybees tend to steer away from regions of the visual field that experience rapid image motion (Srinivasan 2011), indicating that this is a strategy for sensing and avoiding dangerously close objects. When a bee flies through a narrow passage, it positions itself such that both eyes experience approximately the same image velocity (Fig. 1b). This ensures that the two walls of the passage are at the same distance from the bee, enabling a collision-free flight through the middle of the tunnel (Serres et al. 2008, Dyhr and Higgins 2010, Srinivasan 2011). That bees indeed use this strategy has been confirmed by moving the visual texture on one of the walls, either in or against the direction of the bee's flight. The bees then no longer fly through the middle of the tunnel, but along a new axis positioned such that the two eyes again experience equal image velocities (Srinivasan 2011). When an object is approached, its image expands in the eye and the rate of expansion increases as the distance to the object decreases. In the fruit fly, image expansion appears to be a potential cue for detecting nearby objects and avoiding collisions (Tammero and Dickinson 2002).

2.3 Regulating flight speed

Insects appear to regulate their flight speed by measuring and holding constant the average image velocity that is experienced by the eyes. When flying through a tunnel (as in Fig. 1b), bees fly at a speed such that the angular velocities of the images in the lateral fields of view are maintained at a value of about 300 deg/sec (Srinivasan 2011). They fly faster when the patterns on the two walls are moved in the same direction as the bee's flight, and slower when the patterns are moved in the opposite direction, always maintaining the lateral image velocity at the target value (Baird et al. 2005). Fruit flies appear to use the same strategy (Fry et al. 2009). A consequence of using image speed to control the speed of flight is that the flight speed would automatically be reduced when negotiating a narrow passage, and increased when flying in a wide-open environment. Bees indeed display this behaviour when flying through a tapered tunnel (Fig. 1c), or in tunnels of different widths (Srinivasan et al. 1996, Barron and Srinivasan 2006, Baird et al. 2010). Thus, insects appear to have a simple, elegant strategy for automatically reducing the speed of flight to lower and safer levels when flying through densely cluttered environments.

2.4 Regulating flight altitude

When an insect flies at a fixed speed, the velocity of the image of the ground will provide an indication of the insect's altitude – the lower the altitude, the higher the image speed. In principle, therefore, the velocity of the image in the ventral field of view can be used to control flight altitude. Observations of bees flying in a tunnel in which the image speed generated by the floor is artificially manipulated, suggest that bees indeed use this cue (Baird et al. 2006). Movement of the floor pattern in the direction of flight causes the bees to lower their altitude, whilst maintaining the same air speed (Portelli et al. 2010a). In this study the bees maintained a ventral image velocity of about 265 deg/sec, a value not very different from that used to regulate flight speed (ca. 300 deg/sec, see above). *Drosphila* appears to regulate height by sensing expansions or contractions of the image of the ground to increase or decrease altitude accordingly (Straw et al. 2010).

2.5 Orchestrating smooth landings

In order to perform a smooth landing, it is necessary to control flight such that the speed of approach decreases progressively as the ground is approached, to reach a value close to zero just before touchdown. Video films of bees performing grazing landings on horizontal surfaces have shown that the speed of flight is progressively reduced as the surface is approached, in such a way as to hold constant the angular speed of the image of the ground in the eye (Srinivasan et al. 2000b). Holding the image speed constant automatically ensures that flight speed decreases steadily as the ground is approached, leading to a perfectly smooth touchdown (Fig. 1d). The simplicity and elegance of this landing strategy lies in the fact that it does not require knowledge of the instantaneous height above the ground, or the instantaneous speed of flight – all that is required is the measurement (and regulation) of the speed of the image in the eye. A mathematical model of this landing strategy successfully predicts all of the observed characteristics of the landing trajectories (Srinivasan et al. 2000b). During landing, bees maintain image velocities of about 500 deg/s in the ventral field of view (Srinivasan et al. 2000b).

2.6 Gauging distance travelled

Insects that make repeated foraging excursions and return home unerringly must be capable of estimating the distances that they have travelled. A reliable "odometer" would not only help them navigate back home from the food source, but return to it if necessary. Behavioural experiments strongly suggest that bees estimate distance flown by summing (integrating) the optic flow, over time, that is generated by the movement of the image of the environment in their eyes as they fly from the hive to the food source (Srinivasan 2011). Three lines of evidence support this thesis. First, bees report a smaller travel distance in their waggle dances when flying at a high altitude, compared to when they fly at normal cruising heights of 1–2 m above the ground (Esch and Burns 1995, 1996). This is consistent with the fact that the ground generates a lower image velocity when it is further away from the bee. Second, bees that have flown through a short, narrow tunnel report an abnormally large travel distance, compared with bees that have flown a comparable distance in a natural outdoor environment. Evidently, the close proximity of the walls and floor of the tunnel greatly amplify the magnitude of the optic flow that is experienced in this environment, compared to flight in natural environments where surfaces and objects are much further away (Srinivasan et al. 2000a, Esch et al. 2001).Third, when bees are trained to fly to a feeder positioned at a fixed distance into a tunnel and then tested in a fresh tunnel with the feeder removed, they search at the correct location when the test tunnel has the same width as the training tunnel, but at a lower distance when the test tunnel is narrower, and at a higher distance when the test tunnel is wider (Srinivasan et al. 1997b, Hrncir et al. 2003). This, again, is consistent with the fact that narrower tunnels generate a greater magnitude of optic flow, leading to the perception of a greater travel distance. The advantages and disadvantages of measuring distance travelled by using optic flow signals – rather than by measuring energy consumption or time of flight, or by counting wing beats, for example – are reviewed in Srinivasan (2011). Desert ants, on the other hand, gauge travel distance by monitoring the motion of their legs, in a process akin to counting footsteps (Wittlinger et al. 2007).

2.7 The importance of image velocity in controlling and regulating flight

The observation that bees strive to balance the image velocities in the two eyes whilst negotiating narrow gaps, as well as maintain a constant image velocity while regulating flight speed, while controlling altitude and while guiding landing, suggests that measuring and controlling the speed of the image in the eye is a fundamental and overarching principle of visual guidance in insects. That the regulated image speed is in the vicinity of 300–500 deg per sec in all of these behaviours suggests the involvement of the same, or very similar movement-detecting circuitry in each case, and a parsimonious, yet effective utilization of the restricted computational capacity in these miniature brains to deal with a number of tasks. Interneurons in the ventral nerve cord of honeybees, which respond to large field motion stimuli (Ibbotson 2001), and other neurons in the optic lobe that appear to encode image velocity (Straw et al. 2008) may be components of the neural circuits that are involved in guiding all of these behaviours.

3. Applications to robotics

In recent years there has been considerable interest in implementing some of the insights gained from the study of vision and navigation in insects to the guidance of terrestrial and aerial vehicles. We have seen that insects rely heavily on cues derived from optic flow to gauge the distances to various obstacles and surfaces, and to manoeuvre safely in their three-dimensional environment. From the standpoint of machine vision, the use of optic flow cues for navigation is far less demanding, computationally, than is the classical approach of using stereo vision for this purpose. This makes the optic-flow-based approach attractive for implementation in small, compact, and lightweight aerial vehicles (Floreano et al. 2009). The following sections outline some of the ways in which this approach is being used for the guidance of autonomously navigating robots.

3.1 Panoramic imaging using synthetic compound eyes

Compound eyes have the advantages of being low in volume and weight, and endowing the animal with nearly all-round vision (Lee and Szema 2005). One approach to producing a synthetic compound eye has been to use microlithography to create a planar, 2-D array of microlenses, stacked on top of a specially designed CCD chip that allocates one pixel per microlens to create a planar array of synthetic ommatidia (Bruckner et al. 2009). By making the inter-pixel separation slightly smaller than the inter-lens separation, an imaging array is created in which the optical axes of the individual ommatidia are perpendicular to the plane near the centre of the array, but diverge increasingly outwards toward the periphery (Bruckner et al. 2009). While this does not provide a very large field of view (±30 deg being typical), the device is compact (typically 10 mm × 10 mm × 3 mm) and light. Panoramic vision can be achieved, in principle, by using a few of these arrays to cover the entire visual field. Another approach has been to use one photodiode for each ommatidium, and to arrange the photodiodes on an appropriately tessellated sphere (Maddern and Wyeth 2010), or to fabricate a 2-D array of photodiodes on an elastomeric sheet, which is then stretched on a spherical surface to take the shape of a hemisphere (Ko et al. 2008). Yet another technique has produced a compound eye with dimensions and optical characteristics very similar to that in a honeybee, using reconfigurable templating of a 2-D array of microlenses, and ultraviolet (UV) light to form light guides similar to those that are present beneath each ommatidial lens in the real compound eye (Jeong et al. 2006).

3.2 Panoramic imaging using reflective surfaces

A completely different approach to panoramic imaging makes use of a standard video camera in conjunction with a specially shaped, curved reflective surface to increase the field of view of the camera. This is a convenient and flexible approach, because the camera can be purchased off the shelf and the profile of the mirror can be tailored to achieve a variety of objectives. A disadvantage of this approach, however, is that it cannot be designed to be as compact or lightweight as the synthetic compound eyes described above. However, it is easily implemented, does not require exotic fabrication technology, and has already been used extensively in robotic applications. Some examples of these camera-mirror systems are described below.

One system uses the so-called “fixed-

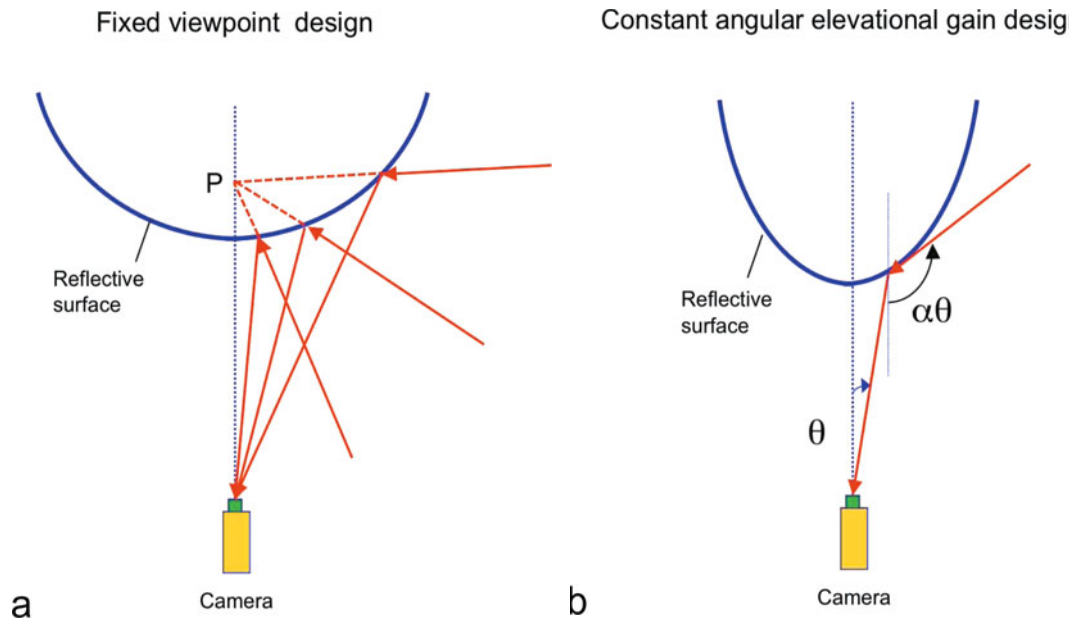


Fig. 2 Two examples of panoramic imaging systems that use a standard video camera in conjunction with a specially shaped reflective surface. **a** Fixed viewpoint system. **b** Constant elevational gain system

viewpoint" design (Fig. 2a). Here the profile of the mirror is designed such that all of the incoming rays that are reflected by the mirror into the camera intersect at a common point on the axis of the mirror, when they are extrapolated as shown by the dotted lines. The image captured by the camera then corresponds to a view of the world from a fixed point P. The resulting mirror surface has the profile of a rectangular hyperbola. The details and design of such a mirror are described in Yagi et al. (1995).

Another system achieves constant angular elevational gain (Fig. 2b). Here the profile of the mirror is such that a ray entering the camera at an angle θ with respect to the optical axis, arrives from the environment at an elevational angle αθ before striking the mirror (Chahl and Srinivasan 1997). The elevational gain α of the mirror can be selected to suit the particular application. A mirror with α = 5, for example, will enable a camera with a nominal viewfield of ±30 deg to have an extended viewfield of ±150 deg in the elevational plane and an all-round view (360 deg) in the azimuthal plane. In this design, equal displacements along the radius of the image correspond to equal changes in the elevational angle of view of the environment (Chahl and Srinivasan 1997). As a result, the camera's pixels sample the environment uniformly in the elevational plane, and are therefore utilized optimally. We note, in passing, that a spherical mirror does not achieve this objective: it compresses the environment near the periphery of the image, and stretches it near the middle. Figure 3a shows an image of the environment, as captured by a constant elevational gain mirror. Figure 3b shows a digitally remapped version of this image, where the vertical axis represents radial distance from the centre of the image (r) and the horizontal axis represents the angular orientation ϕ of the radius. Equal increments in r in the image of Fig. 3a correspond to equal vertical

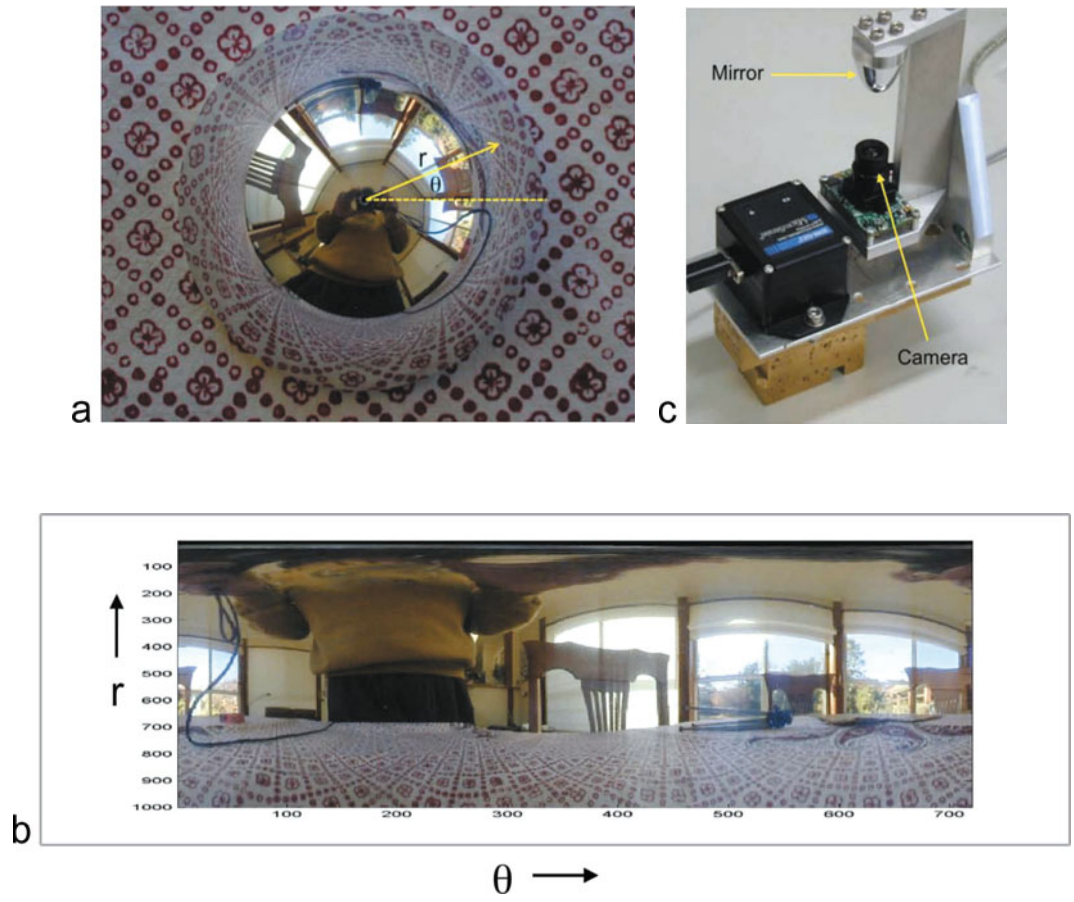


Fig. 3 Constant elevational gain mirror showing **a** image of the surrounding environment as reflected by the mirror and captured by an overhead camera. **b** A digitally remapped version of **a** in which the vertical axis represents radius and the horizontal axis represents the angular orientation of this radius. **c** An example of such a system, using a mirror of diameter 21 mm and depth 20 mm, installed on a model aircraft

displacements in the image of Fig. 3b, and to equal changes in the elevational angle of view of the environment. The details and design of these mirror profiles are described in Chahl and Srinivasan (1997). Figure 3c shows an example of such a system, installed on a model aircraft.

It should be noted that the constant elevational gain system described above does not possess a fixed viewpoint, and that the fixed viewpoint system described above does not have a constant elevational gain. However, a system that achieves both objectives can be created by using two reflective surfaces, or one reflective and one refractive surface (Stuerzl and Srinivasan 2010).

Another mirror-based design, useful in measuring the large magnitudes of optic flow that are generated by the ground during low-altitude flight at high speeds, uses a camera and a specially shaped convex mirror profile as shown in Fig. 4a, b to achieve two objectives. They are: (i) To reduce the speed of the image of the ground to values that are low enough to permit accurate measurement; and (ii) to remove the per-

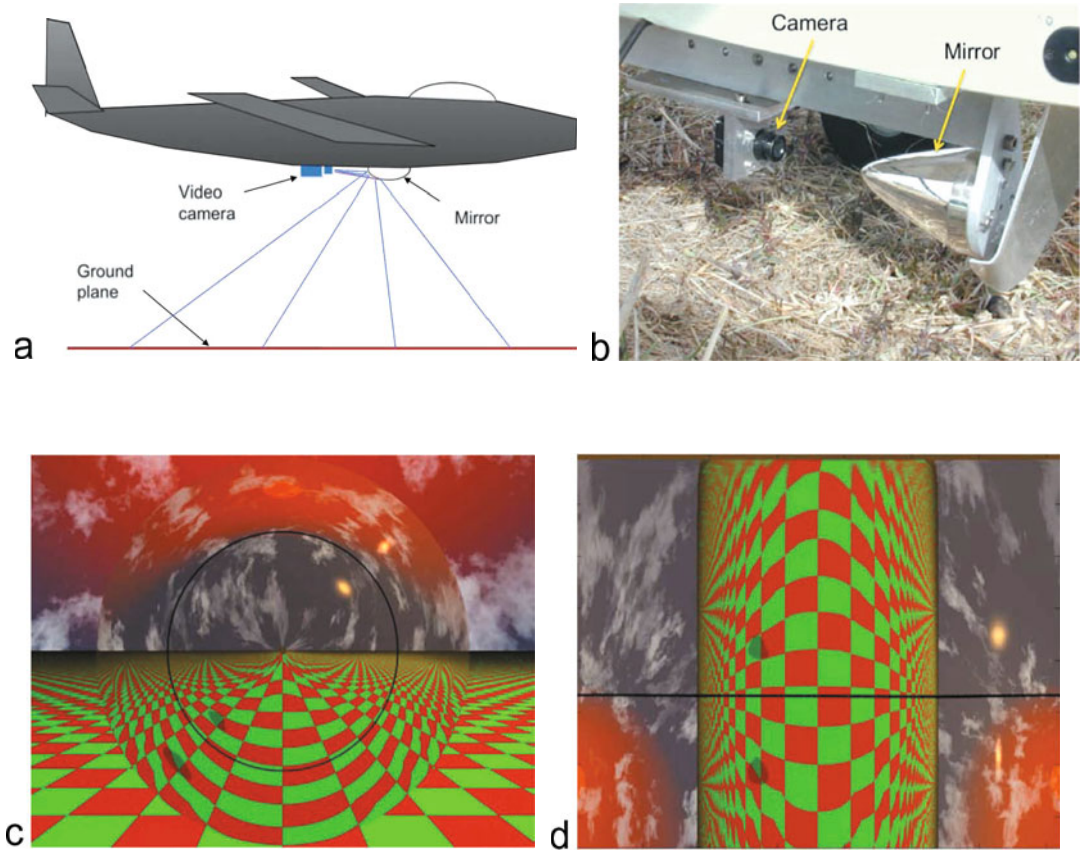


Fig. 4 Illustration of vision system for terrain following. **a** A video camera views the ground through a specially shaped, curved reflecting surface. **b** View of system installed on the underside of a model aircraft. **c** Simulation of view in the mirror as captured by the camera during horizontal flight over an infinite horizontal plane. **d** Unwarped version of **c**, remapped as in Fig. 3b, showing the removal of the perspective distortion along the vertical axis. **a**, **c** and **d** are adapted with permission from Srinivasan et al. (2006)

spective-induced distortion of different regions of the ground in front of the aircraft (see Fig. 4c). This ensures that there is no variation in image velocity along the vertical axis of the remapped image (Fig. 4d), thus promoting the accurate measurement of optic flow along this axis (Srinivasan et al. 2006). This way of mapping the world is similar in some respects to that achieved by the compound eyes of semi-terrestrial crabs that live on a flat substrate (Zeil et al. 1986) – although the objective in the case of the crab seems to be translate range measurements into angular measurements in the eye, rather than to measure image velocities. An additional advantage of imaging the world in this way is that the largest magnitude of the optic flow that is sensed by the system defines the radius of a "collision-free" cylinder of space (a clear zone) through which the aircraft can fly safely (Fig. 5a). The smaller the flow magnitude, the larger the radius of this cylinder, and the safer the passage.

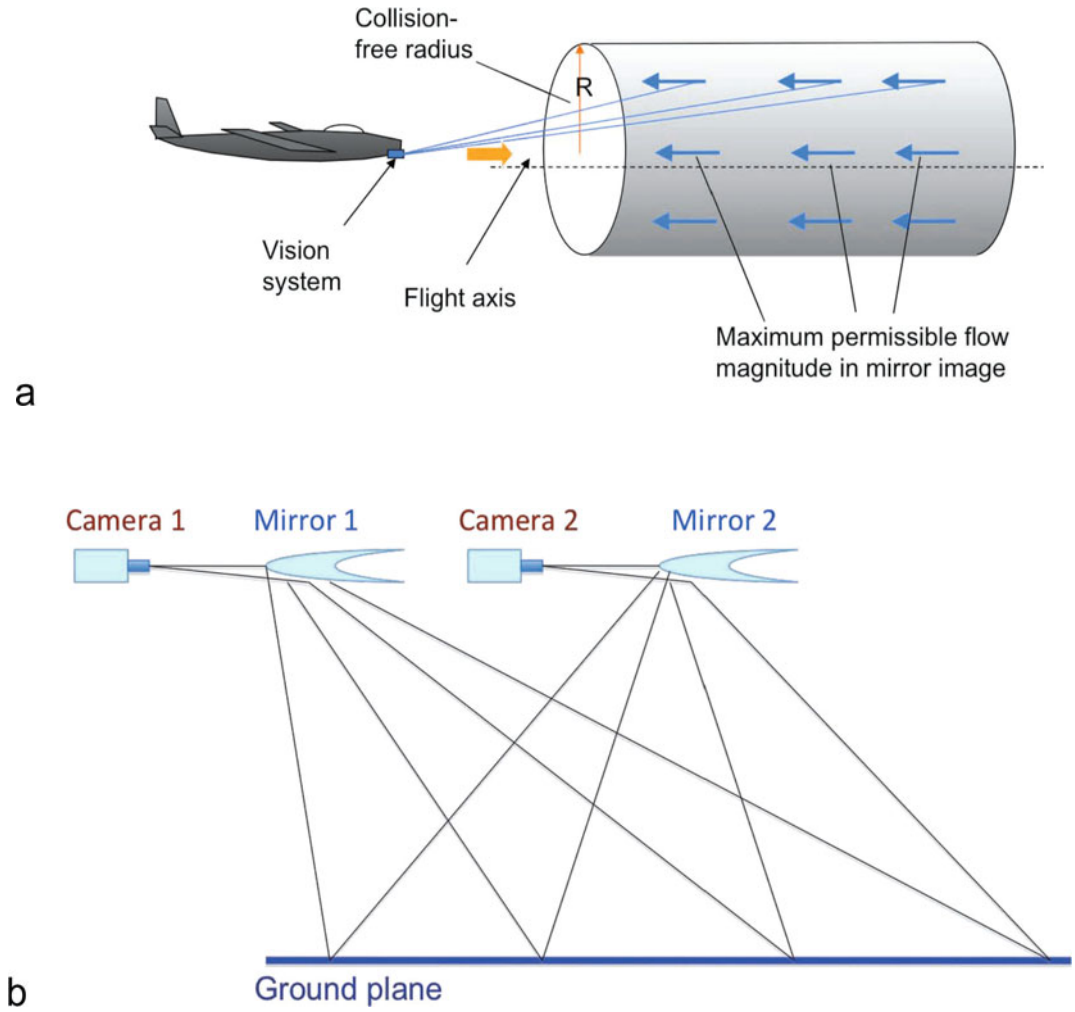


Fig. 5 **a** Illustration of how the terrain-following vision system of Fig. 4 defines a collision-free cylinder whose radius depends upon the highest magnitude of optic flow that is measured. **b** Illustration of twin coaxial camera/mirror system for measuring virtual translational optic flow. **a** is adapted with permission from Srinivasan et al. (2006) and **b** from Moore et al. (2009)

3.3 Guidance of robots along corridors

The principle by which bees fly safely through narrow passages, offers a simple strategy for guiding a robot along a corridor. By balancing the speeds of the images of the two side walls, the robot can be steered along the middle of the corridor without colliding with the walls. Furthermore, the speed of the robot can be adjusted to a safe value by holding constant the average velocity of the images of the two walls. The feasibility of this technique has been demonstrated many times over the past 15 years, in simulations (e. g. Portelli et al. 2010 b) as well as in real, land-based robots (e. g. Humbert and Hyslop 2010, Srinivasan et al. 2010, Srinivasan 2011). Algorithms for measuring image motion are described in (Srinivasan 1990, 1994; Srinivasan et al. 1997 a).

3.4 Terrain following guidance for aircraft

Flying at a constant, low height above the ground is important when there is a need to perform close-up photographic exploration of terrain, or, in a military application, to evade detection by enemy radar. If the ground speed of the aircraft is known (e. g. through measurement of airspeed, or from GPS information), then, following the example of the honeybee, the height above the ground can be computed and regulated by measuring the optic flow that is generated by the image of the ground. The optic-flow based approach is attractive because it only requires the presence of a small, inexpensive, low-resolution video camera on board. This strategy of terrain following has been implemented successfully in fixed-wing (Barrows et al. 2003) as well as rotary-wing aircraft (Garratt and Chahl 2008) and has several advantages over traditional methods of height measurement that use heavy, bulky and power hungry instrumentation such as radar or ultrasound.

When using optic flow to compute the range of an object, it is important to bear in mind that it is only the *translation-induced* component of optic flow that contains information on range – nearer objects induce higher magnitudes of optic flow. The flows induced by any accompanying rotations of the aircraft – for example, yaw, pitch or roll – do not carry information on range, because their magnitudes are independent of range. Consequently, we cannot work directly with the raw optic flow readings. It is necessary to (i) measure the rates of yaw, roll and pitch using gyroscopes, to (ii) use the gyroscopic readings in conjunction with pre-computed optic-flow templates for yaw, roll and pitch to determine the rotation-induced flow components, and to (iii) calculate the translation-induced component of optic flow by subtracting, or "peeling off" the flow components that are generated by the rotations. The residual pattern of optic flow, which is purely that induced by the translational component of the aircraft's motion, can then be used compute range. Details of this procedure are described in (Soccol et al. 2007).

A simpler solution, however, would be to artificially create the effect of a purely translatory forward motion of the aircraft by configuring *two* of the camera/mirror systems of Fig. 4 as shown in Fig. 5 b. The apparent optic flow that is measured by comparing the images acquired by these two systems *at the same instant of time* is then exactly equal to the flow that would have been registered by the rear mirror system if it (and the aircraft) had made a pure translation to the position of the front mirror system (Moore et al. 2009). The "virtual optic flow" generated by this coaxially arranged, dual camera-mirror system provides an accurate and immediate reproduction of the translation-induced component of the flow, without having to move the aircraft at all. This flow can then be used for a variety of purposes, including (i) computing a range map of the environment, (ii) fitting a plane to the range data to estimate the distance and the orientation of the ground plane relative to the aircraft, and (iii) using this information in a feedback loop to control the aircraft's altitude and attitude (Moore et al. 2011).

Figure 6 a shows an example of a vision system that uses virtual optic flow to control a model aircraft, a Super Frontier Senior-46 (wingspan 2040 mm), in which the engine and propeller have been repositioned above the wing to make room for the vision system (Fig. 6 b) (Moore et al. 2010). Figure 6 c shows the performance of the system in regulating altitude and attitude, where control alternates between the *manual* mode (when the aircraft is controlled by a pilot) and the *auto* mode (when the autopilot is in operation). In the *auto* mode the system is commanded to hold the altitude at 10 m and

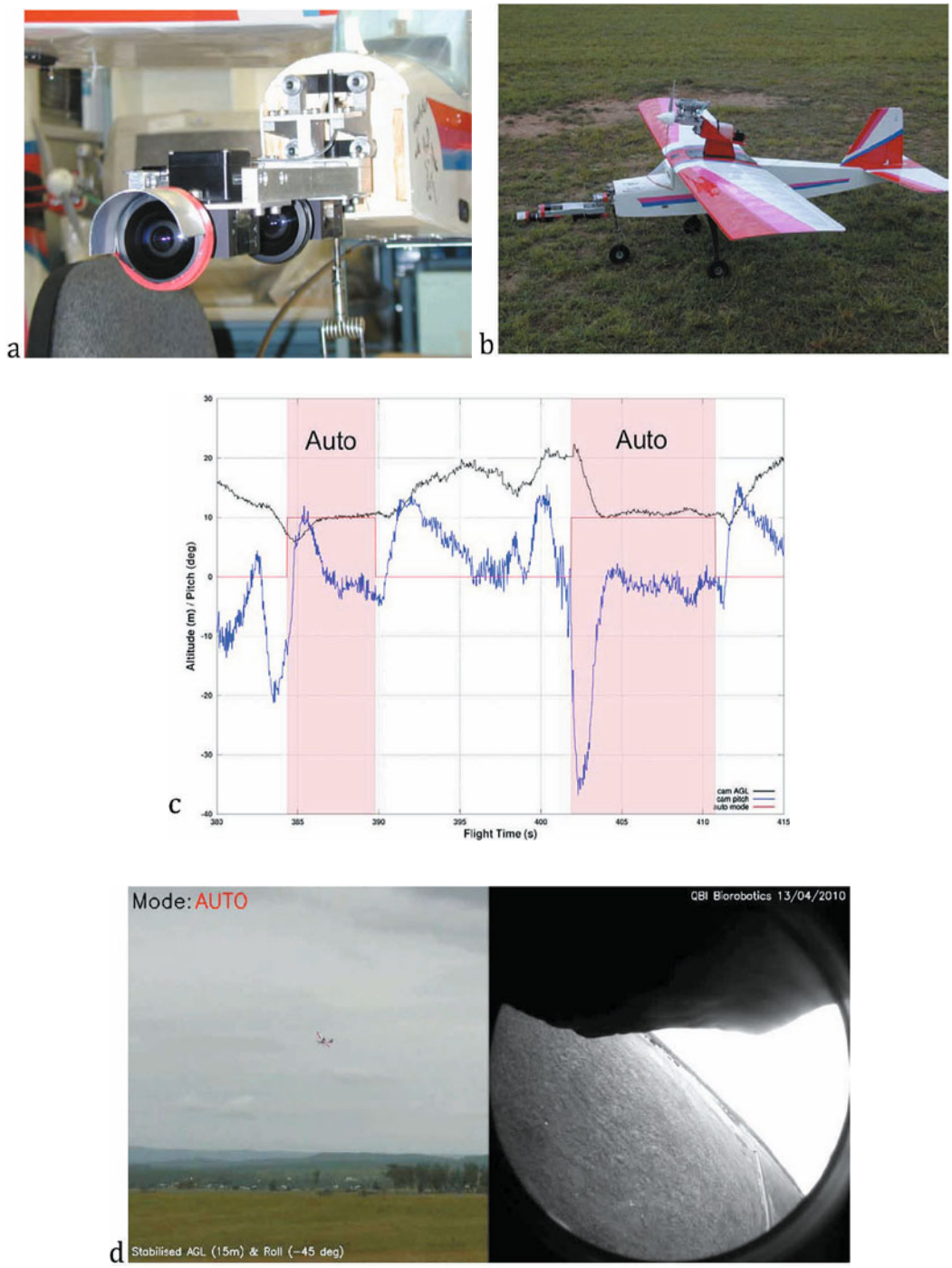


Fig. 6 Vision system for automatic control of aircraft altitude and attitude. **a** View of the two coaxial cameras. **b** View of vision system mounted on aircraft. **c** System performance, showing altitude (black trace) and pitch (blue trace) during periods of manual flight (white background) and automatic flight (pink background). During the *auto* mode the system is commanded to hold the aircraft at an altitude of 10 m and a pitch of 0 deg. **d** View of aircraft (left panel) and image acquired by the front camera (right panel) during an *auto* mode in which the aircraft is commanded to maintain an altitude of 15 m and a roll angle of –45 deg

the pitch at 0 deg. It is evident from Fig. 6c that the system performs these functions well. When control is passed from *manual* to *auto* at a height of 20 m and a pitch of –35 deg, the system is able to attain the target parameters within about 3 seconds. The left-hand panel of Fig. 6d is a view of the aircraft flying in a different *auto* mode where it is commanded to maintain an altitude of 15 m and a roll angle of –45 deg. The right-hand panel shows an image captured by the front camera at the same instant of time. References to other work on visually-guided terrain following can be found in Moore et al. (2011), Srinivasan et al. (2009), and Srinivasan (2011).

3.5 Control of aircraft landing

We have seen that honeybees perform grazing landings by adjusting their flight speed so as to hold the image velocity of the ground constant during the approach. Autonomous landing approaches using approximations of this strategy have been achieved with fixed wing aircraft (Chahl et al.2004, Beyeler 2009). It is difficult to implement the honeybee landing strategy exactly on a fixed-wing aircraft, because this requires the ground speed to approach arbitrarily small values (below stall speed) as the aircraft nears the ground. Consequently, a modified landing strategy has been implemented and tested, in which the throttle is cut and the elevator setting is adjusted in closed loop to hold the magnitude of the optic flow from the ground constant as the altitude drops. The result is that altitude and forward speed decrease approximately linearly with time as the ground is approached, ensuring a smooth touchdown (Chahl et al. 2004, Beyeler 2009).

3.6 Vision-based odometry

Visual odometry, using the honeybee-inspired principle of integrating optic flow, has been implemented successfully in a number of terrestrial robots. In one example a robot was able to traverse a corridor repeatedly, always stopping after it had travelled a fixed distance as reported by its visually driven odometer. The error in the stopping position was less than 2 % of the traverse distance (Weber et al. 1997). In another study (Chahl and Srinivasan 1996), a robot measured optic flow to compute all of its translations and rotations along any arbitrary route that it took from a starting point, and was able to use this information to (a) compute where it was in relation to the starting point, at all times; and (b) to return successfully to the starting location – again relying solely on optic flow information, and without knowledge of or reference to any landmarks in the environment. Visual odometry based on optic flow has also been implemented successfully in large automotive vehicles, by using a downward-facing video camera to view the road surface (Nourani-Vatani et al. 2009).

3.7 Horizon-based control of aircraft attitude

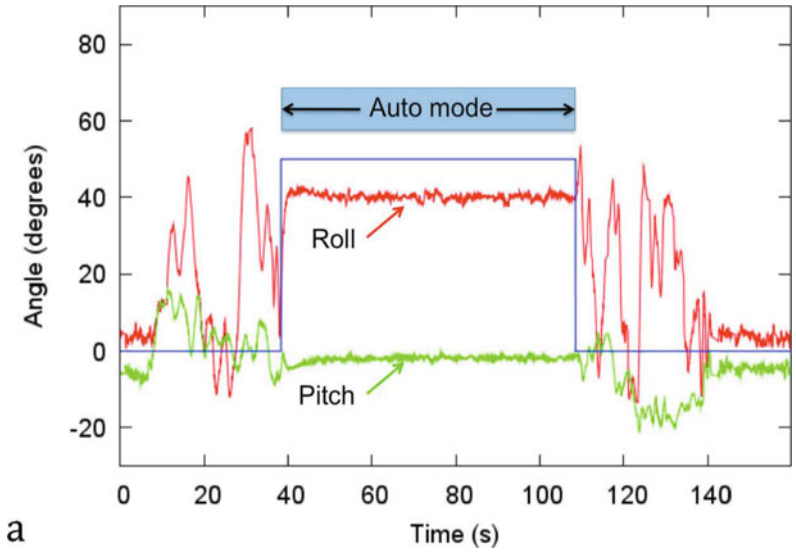
Conventional methods of estimating and stabilizing attitude involve the use of gyroscopes to measure the rates of roll and pitch. This approach has the disadvantage that the instantaneous attitude is determined by integrating these rates over time, which means that the error in the estimate of the attitude will increase with time, due to the presence of noise in the rate signals. The problem does not arise if the horizon is used to estimate attitude, because this visual feature provides an absolute external reference and there is no cumulative error arising from integrating angular rates over time. This is

precisely what insect ocelli seem to be organized to achieve.

To date, two kinds of ocellus-inspired approaches have been explored to control aircraft attitude. One approach uses three photodiodes to play the roles of the ocelli. When the aircraft rolls counterclockwise, the left photodiode views a smaller patch of the sky than does the right photodiode, and therefore produces a smaller signal. The resulting error signal is used to generate a corrective roll command through the ailerons. The presence of the sun in the sky can cause a serious imbalance in the signals from the two lateral photodetectors if one of them happens to be viewing the sun. This problem can be overcome by computing, within each detector, the ratio of the ultraviolet (UV) component to the green component of the incident light. Using this ratio signal offers two advantages. First, the ratio is largely insensitive to presence of the sun, because the spectral composition of the light from the sun is very similar to that from the rest of the sky. Second, the ratio is invariant to the mean level of illumination and enables the system to function over a wide range of ambient light intensities. This may be one of the reasons for the UV-green antagonism that has been observed in interneurons of the ocellar pathway (van Kleef et al. 2005). Pitch is monitored by the median ocellus, which measures light within a broad visual field directed at the frontal horizon. A level pitch attitude is maintained by comparing the UV-to-green ratio signal from the median ocellus with the average of the ratio signals from the left and right ocelli. The resulting error signal is zero when the horizon bisects the visual field of the median ocellus, irrespective of the roll attitude of the aircraft. It thus provides a robust pitch signal that can be used to apply elevator/rudder control to counteract unwanted pitch disturbances. A hardware implementation of this system in fixed-wing aircraft models has produced very satisfactory results (Chahl et al. 2003).

Another approach, which incorporates the principles of ocellar function but does not use electronic ocelli *per se*, is to detect the location and profile of the horizon in the image captured by one or more onboard cameras. The colour of each pixel in the image is analysed to determine whether it belongs to the ground or to the sky (Todorovic and Nechbya 2004, Thurrowgood et al. 2009). This determination is carried out by generating a signal-derived variable, U, which is an optimally weighted linear combination of the Red, Green and Blue colour components of the pixel. The pixel is declared to belong to the sky or the ground according to whether U is greater or lower than a preset threshold. This procedure enhances the speed of discrimination and maximises its accuracy (Thurrowgood et al. 2009). Horizon pixels are identified as those lying at the transition between ground and sky. The positions of these horizon pixels in the image are then back-projected into the external environment using the known geometry of the camera's optics, and a 3-D plane is fitted to these points as detailed in (Thurrowgood et al. 2009). The orientation of this plane relative to the vision system then specifies the attitude of the aircraft relative to the horizon, thus determining its pitch and roll. This scheme, like the electronic ocellar system, is robust to the position of the sun and to changes in illumination because it locates the horizon by using information on colour, rather than intensity. A further advantage of this system is that it does not require a separate set of optoelectronic sensors for monitoring aircraft attitude – the algorithm can be incorporated into a vision system that is already place, e. g. for sensing optic flow.

The performance of this scheme is illustrated in Fig. 7a, which shows the results of a test in which the aircraft was initially flown manually, and subsequently commanded



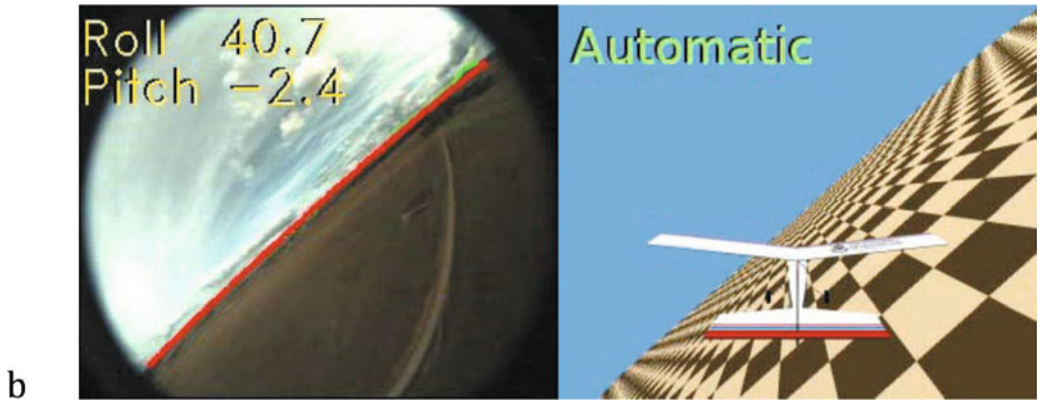


Fig. 7 Insect ocellus-inspired system for measuring and stabilizing aircraft attitude by monitoring the horizon. **a** Illustration of system performance. The graphs show pitch (green trace) and roll (red trace) during the *manual* flight mode (blue trace low), and during the *auto* mode (blue trace high) when the system is commanded to hold a pitch angle of 0 deg and a roll angle of 40 deg. **b** The left-hand panel shows an image captured by the vision system when the aircraft is under automatic control and commanded to maintain the pitch and roll angles indicated above. The right-hand panel shows an animation of the orientation of the aircraft relative to the ground plane at the same instant of time

hold a pitch angle of 0 deg and a roll angle of 40 deg during the time interval 38–107 sec. Manual control was resumed after 107 sec. It is evident that the aircraft maintains the prescribed attitude in a stable and accurate manner during the period of automatic control. The left-hand panel of Fig. 7b shows one frame captured by the camera during the *auto* mode, with the red line showing the estimated horizon profile and the numbers indicating the computed values of roll and pitch. The right-hand panel of Fig. 7b shows an animated reconstruction of the instantaneous orientation of the aircraft relative to the ground plane. Recent work has demonstrated that this technique for monitoring and controlling aircraft attitude can be used to automate the execution of a variety of

aerobatic manoeuvres such as loops and Immelmann turns (Thurrowgood et al. 2010).

Conclusions

In performing tasks that require sophisticated visuomotor co-ordination, it appears that insects are unhampered by their miniscule brains and (presumably) limited processing capacity. Instead, they seem to have evolved simple, elegant and alternative solutions to deal with many of their everyday tasks. In order to overcome the limitations imposed by their small interocular separations, many insects appear to rely heavily on cues derived from optic flow – rather than stereo vision – to avoid collisions with objects, to control flight speed, to guide landings, and to perform odometry. It appears that many of these biological principles of visual guidance can be applied usefully to the design of effective and efficient algorithms for the guidance of airborne vehicles.

Outstanding Questions

Many enigmas remain, some of which are listed below.

The neural basis of movement detection While the neurons in the visual pathways underlying the optomotor behavior are now beginning to be well characterized and understood – at least in flies – there are still major gaps. For example, we still do not know exactly where along the visual pathway the fundamental operation of detecting the movement of the image and sensing its direction takes place, although suggestions have been made and models proposed (rev. Borst 2009).

The neural basis of other movement-sensitive behaviors A major puzzle concerns the neural basis of the other movement-sensitive behaviors such as obstacle avoidance, control of flight speed, and visual odometry. All of these behaviors appear to be mediated by movement-detecting systems that measure the speed of the image reliably, and largely independently of the spatial texture of the image, or of its contrast. Robust measurement of image speed is critically important in controlling behaviors such as collision avoidance, flight speed regulation, and visual odometry. This property is not exhibited by the movement-detecting neurons that are presumed to mediate the optomotor response. An important and unresolved question, then, is whether the behaviours that depend on the accurate measurement of image speed are mediated by a different population of movement-sensitive neurons, as yet undiscovered and unexplored.

Detection and pursuit of other objects in the environment Insects such as flies and dragonflies are extremely adept at detecting, pursuing and intercepting other small, moving targets such as prey or mates. While this behaviour has been modelled quantitatively with some success (Land and Collett,1974, Olberg et al. 2000, Mizutani et al. 2003) the neural mechanisms that underlie this behaviour remain largely unexplored, although progress is encouraging (Nordstrom et al. 2006, Geurten et al. 2007). Implementation of these "spot, chase and intercept" strategies in aircraft should be an exciting challenge in the coming years, leading to a variety of applications.

(1) *Gyroscopic organs?* While many insects possess two pairs of wings, in dipteran insects the hind pair have been modified into tiny, club-shaped structures, barely visible to the unaided eye, that oscillate up and down in synchrony with the front wings, but in antiphase. These structures, known as the halteres, are believed to function as gyroscopes which sense the coriolis forces that accompany the insect's rotations about the yaw, pitch and roll axes (Hengstenberg 1993, Dickinson 1999) (see also Chapter IV.19 by T. Daniel et al.). Behavioral experiments with houseflies indicate that the haltere system senses rapid rotations, thus augmenting the contribution of the visual system – which responds only to slower rotations – in stabilizing attitude. Since the haltere system does not rely on vision, it can, in principle, help stabilize attitude even during flight in the dark or in a featureless environment. Honeybees and other hymenopteran flying insects, on the other hand, possess two pairs of wings. This raises a number of questions: Do honeybees lack a gyroscopic sense, given that they do not possess halteres? Do the hind wings provide a haltere-like function in these insects? Or are there other organs or structures that serve this purpose? In the hawkmoth, for example, it has been suggested that the antennae func-

tion as gyroscopic sensors (Sane et al. 2007), in addition to playing an important role in olfactory, anemometric and tactile perception.

A magnetic compass? Equally intriguing is the emerging possibility that honeybees use a magnetic sense to aid their orientation and navigation. Early experiments showed that honeybees can be trained, by reward in a dual-choice paradigm, to distinguish between different patterns of magnetic fields (Walker and Bitterman 1985, Kirschvink and Kobayashi 1991). Bees also appear to be able to sense the earth's magnetic field and use it as an earth-based orientational reference while learning the visual environment around a newly discovered food source (Collett and Baron 1994). Trophocyte cells in honeybees contain granules of superparamagnetic magnetite, and these cells receive sensory innervation (e.g. (Hsu et al. 2007)). Thus, there is some evidence for the existence of a magnetic sense in the honeybee, and for a plausible neural substrate. The next steps will be to (i) design further behavioural experiments to better understand the contexts in which bees sense and use the earth's magnetic field, to (ii) investigate the physiological mechanisms that underlie transduction of the magnetic field, to (iii) examine how the direction of the geomagnetic field is represented in the nervous system, and to (iv) investigate whether, and, if so, how this information is combined with information from the bee's sun and polarization compasses (which are reviewed, for example, in Wehner and Labhart 2006) to mediate navigation.

Outlook

Technology has now advanced to the point where it is nearly possible to (a) place a tethered, flying insect within a high-definition, panoramic, colour visual display (e.g. (Luu et al., in press)); (b) measure the forces (thrust and lift), and torques (yaw, roll, and pitch) exerted by the insect during its tethered flight, and (c) use these signals to control the visual display to reproduce the consequences of the insect's intended movements (Straw et al. 2010) (Taylor et al. 2011). Such a virtual-reality system would simulate free flight and permit a detailed, quantitative, and tightly controlled investigation of the visuomotor mechanisms that are involved in controlling insect flight. Another approach, on the threshold of technological feasibility, is the prospect of (a) mounting miniature accelerometers and gyroscopes on freely flying insects to record and transmit their motions wirelessly to a base station; and (b) simultaneously recording and transmitting the electrophysiological responses of neurons in the visuomotor pathways (e.g. (Olberg and Leonardo 2010)). This should provide a wealth of new information about the processing of sensory information and the control of flight under natural conditions.

Finally, we should bear in mind that in designing biologically inspired machines and robots, it may not always be beneficial to duplicate the living organism in every possible detail, because animals may not need to address exactly the same set of requirements as a robot. The solution that a particular animal has evolved could be a compromise that attempts to simultaneously fulfil a number of (possibly disparate) goals, many of which may not be relevant to the intended engineering application. A more fruitful approach, therefore, might be to design systems that embody some of the *principles* by which animals acquire, process and use sensory information, rather than to build machines that are slavishly biomimetic. This may be the most fruitful way in which engineering can draw inspiration from biology.

Acknowledgements

Some of the research described here was supported partly by the Australian Research Council Centre of Excellence in Vision Science (CE0561903), a Queensland Smart State Premier's Fellowship, US Army Research Office grant MURI ARMY-W911NF041076, grants N00014-99-1-0506 and N66001-00-C-8025 from the U.S. Defence Advanced Research Projects Agency, and US ONR Award N00014-04-1-0334.

References

Baird E, Kornfeldt T, Dacke M (2010) Minimum viewing angle for visually guided ground speed control in bumblebees. J Exp Biol 213: 1625–1632

Baird E, Srinivasan MV, Zhang SW, Cowling A (2005) Visual control of flight speed in honeybees. J Exp Biol 208: 3895–3905

Baird E, Srinivasan MV, Zhang SW, Lamont R, Cowling A (2006) Visual control of flight speed and height in the honeybee. From Animals to Animats 9, Proc 4095: 40–51

Barron A, Srinivasan MV (2006) Visual regulation of ground speed and headwind compensation in freely flying honey bees (*Apis mellifera* L.). J Exp Biol 209: 978–984

Barrows GL, Chahl JS, Srinivasan MV (2003) Biologically inspired visual sensing and flight Control. Aeronautical J 107: 159–168

Beyeler A (2009) Vision-based control of near-obstacle flight. Ecole Polytechnique Federale de Lausanne, Lausanne

Borst A (2009) *Drosophila*'s view on insect vision. Curr Biol 19: R36-R47

Bruckner A, Duparre J, Wippermann F, Dannberg P, Brauer A (2009) Microoptical artificial compound eyes. In: Floreano D, Zufferey JC, Srinivasan MV, Ellington C (eds) Flying insects and robots. Springer-Verlag, Berlin, Heidelberg, pp 127–142

Chahl J, Thakoor S, Le Bouffant N, Stange G, Srinivasan MV, Hine B, Zornetzer S (2003) Bioinspired engineering of exploration systems: A horizon sensor/attitude reference system based on the dragonfly ocelli for mars exploration applications. J Robotic Systems 20: 35–42

Chahl JS, Srinivasan MV (1996) Visual computation of egomotion using an image interpolation technique. Biol Cybern 74: 405–411

Chahl JS, Srinivasan MV (1997) Reflective surfaces for panoramic imaging. Applied Optics 36: 8275–8285

Chahl JS, Srinivasan MV, Zhang SW (2004) Landing strategies in honeybees and applications to uninhabited airborne vehicles. Int J Robot Res 23: 101–110

Collett TS, Baron J (1994) Biological compasses and the coordinate frame of landmark memories in honeybees. Nature 368: 137–140

Dickinson MH (1999) Haltere-mediated equilibrium reflexes of the fruit fly, *Drosophila melanogaster.* Phil Trans R Soc B 354: 903–916

Dyhr JP, Higgins CM (2010) The spatial-frequency tuning of optic-flow-dependent behaviors in the bumblebee *Bombus impatiens*. J Exp Biol 213: 1643–1650

Egelhaaf M (2008) Fly vision: Neural mechanisms of motion computation. Curr Biol 18: R339–R341

Esch H, Burns J (1996) Distance estimation by foraging honeybees. J Exp Biol 199: 155–162

Esch HE, Burns JE (1995) Honeybees use optic flow to measure the distance of a food source. Naturwissenschaften 82: 38–40

Esch HE, Zhang SW, Srinivasan MV, Tautz J (2001) Honeybee dances communicate distances measured by optic flow. Nature 411: 581–583

Floreano D, Zufferey J-C, Srinivasan MV, Ellington C (2009) Flying insects and robots. Springer, Berlin, Heidelberg

Fry SN, Rohrseitz N, Straw AD, Dickinson MH (2009) Visual control of flight speed in Drosophila melanogaster. J Exp Biol 212: 1120–1130

Garratt MA, Chahl JS (2008) Vision-based terrain following for an unmanned aircraft. J Field Robotics 25: 284–301

Geurten BRH, Nordstrom K, Sprayberry JDH, Bolzon DM, O'Carroll DC (2007) Neural mechanisms underlying target detection in a dragonfly centrifugal neuron. J Exp Biol 210: 3277–3284

Goodman L (2003) Form and function in the honeybee. Internat Bee Res Assoc, Cardiff, U. K.

Hengstenberg R (1993) Multisensory control in insect oculomotor control systems. In: Miles FA, Wallman J (eds) Visual motion and its role in the stabilization of gaze. Elsevier, Amsterdam, pp 285–298

Hrncir M, Jarau S, Zucchi R, Barth FG (2003) A stingless bee (*Melipona seminigra*) uses optic flow to estimate flight distances. J Comp Physiol A 189: 761–768

Hsu CY, Ko FY, Li CW, Fann K, Lue JT (2007) Magnetoreception system in honeybees (*Apis mellifera*). PLoS One 2: e395

Humbert JS, Hyslop AM (2010) Bioinspired visuomotor convergence. Ieee T Robot 26: 121–130

Ibbotson MR (2001) Evidence for velocity-tuned motion-sensitive descending neurons in the honeybee. Proc R Soc Lond B 268: 2195–2201

Jeong K-H, Kim JH, Lee LP (2006) Biologically inspired artificial compound eyes. Science 312: 557–561

Joesch M, Plett J, Borst A, Reiff DF (2008) Response properties of motion-sensitive visual interneurons in the lobula plate of *Drosophila melanogaster.* Curr Biol 18: 368–374

Kirschvink JL, Kobayashi A (1991) Is geomagnetic sensitivity real? Replication of the Walker-Bit-

terman magnetic conditioning experiment in honey bees. Am Zool 31: 169–185

Ko HC, Stoykovich MP, Song J, Malyarchuk V, Choi WM, Yu C-J, Geddes III JB, Xiao J, Wang S, Huang Y, Rogers JA (2008) A hemispherical elecronic eye camera based on compresible silicoan optoelectronics. Nature 454: 748–753

Krapp HG (2000) Neuronal matched filters for optic flow processing in flying insects. Int Rev Neurobiol 44: 93–120

Land MF, Collett TS (1974) Chasing behaviour of houseflies (*Fania cannicularis*). A description and analysis. J Comp Physiol 89: 331–357

Lee LP, Szema R (2005) Inspirations form biological optics for advanced photonic systems. Science 310: 1148–1150

Luu T, Cheung A, Ball D, Srinivasan MV (in press) Honeybee flight: A novel 'streamlining' response. J Exp Biol

Maddern W, Wyeth G (2010) Egomotion estimation with a biologically-inspired hemispheric camera. 12th Australasian Conf on Robotics and Automation, Brisbane

Mizutani A, Chahl JS, Srinivasan MV (2003) Motion camouflage in dragonflies. Nature 423: 604–604

Moore RJD, Thurrowgood S, Bland D, Soccol D, Srinivasan M (2010) UAV altitude and attitude stabilization using a coaxial stereo vision system. IEEE Internat Conf on Robotics and Automation. IEEE Press, Anchorage, Alaska

Moore RJD, Thurrowgood S, Bland D, Soccol D, Srinivasan MV (2011) A bio-inspired stereo vision system for guidance of autonomous aircraft. In: Bhatti A (ed) Advances in theory and aplications of stereo vision. InTech Publishers, Rijeka

Moore RJD, Thurrrowgood S, Bland D, Soccol D, M. V. S (2009) A stereo vision system for UAV guidance. IEEE/RSJ Internat Conf on Intelligent Robots and Systems, St. Louis, Missouri, USA

Nordstrom K, Barnett PD, Moyer de Miguel I, O'Carroll DC (2008) Sexual dimorphism in the hoverfly motion vision pathway. Curr Biol 18: 661–667

Nordstrom K, Barnett PD, O'Carroll DC (2006) Insect detection of small targets moving in visual clutter. PloS Biology 4: e54

Nourani-Vatani N, Roberts J, Srinivasan MV (2009) Practical visual odometry for car-like vehicles. IEEE Internat Conf on Robotics and Automation. IEEE Press, Kobe, Japan

Olberg RM, Leonardo A (2010) Towards wireless monitoring of neural activity during dragonfly prey interception flights. Ninth Internat Congress of Neuroethology. International Society for Neuroethology, Salamanca, Spain, p 25

Olberg RM, Worthington AH, Venator KR (2000) Prey pursuit and interception in dragonflies. J Comp Physiol A 186: 155–162

Portelli G, Ruffier F, Franceschini N (2010a) Honeybees change their height to restore their optic flow. J Comp Physiol A 196: 307–313

Portelli G, Serres J, Ruffier F, Franceschini N (2010b) Modelling honeybee visual guidance in a 3-D environment. J Physiol-Paris 104: 27–39

Reichardt W (1969) Movement perception in insects. In: Reichardt W (ed) Processing of optical data by organisms and by machines. Academic Press, New York, pp 465–493

Sane SP, Dieudonne A, Willis MA, Daniel TL (2007) Antennal mechanosensors mediate flight control in moths. Science 315: 863–866

Serres JR, Masson GP, Ruffier F, Franceschini N (2008) A bee in the corridor: centering and wall-following. Naturwissenschaften 95: 1181–1187

Soccol D, Thurrowgood S, Srinivasan MV (2007) A vision system for optic-flow-based guidance of UAVs. Proc, Ninth Australasian Conf on Robotics and Automation, Brisbane

Srinivasan M (1993) How insects infer range from visual motion. In: Miles F, Wallman J (eds) Visual motion and its role in the stabilization of gaze. Elsevier, Amsterdam, pp 139–156

Srinivasan M, Thurrowgood S, Soccol D (2009) From flying insects to autonomously navigating robots. IEEE Robotics and Automation Magazine 16: 59–71

Srinivasan M, Zhang S, Lehrer M, Collett T (1996) Honeybee navigation en route to the goal: visual flight control and odometry. J Exp Biol 199: 237–244

Srinivasan MV (1990) Generalized gradient schemes for the measurement of 2-dimensional image motion. Biol Cybern 63: 421–431

Srinivasan MV (1994) An image-interpolation technique for the computation of optic flow and egomotion. Biol Cybern 71: 401–415

Srinivasan MV (2011) Honeybees as a model for the study of visually guided flight, navigation, and biologically inspired robotics. Physiol Rev 91: 389–411

Srinivasan MV (in press) Visual control of naviga-

tion in insects and its relevance for robotics. Curr Opinion Neurobiol

Srinivasan MV, Chahl JS, Nagle MG, Zhang SW (1997a) Embodying natural vision into machines. In: Srinivasan MV, Venkatesh S (eds) From living eyes to seeing machines. Oxford University Press, U. K., pp 249–265

Srinivasan MV, Thurrowgood S, Soccol D (2006) An optical system for guidance of terrain following in UAVs. IEEE Internat Conf on Advanced Video and Signal Based Surveillance (AVSS '06). IEEE Press, Sydney, pp 51–56

Srinivasan MV, Zhang SW, Altwein M, Tautz J (2000a) Honeybee navigation: Nature and calibration of the "odometer". Science 287: 851–853

Srinivasan MV, Zhang SW, Bidwell NJ (1997b) Visually mediated odometry in honeybees. J Exp Biol 200: 2513–2522

Srinivasan MV, Zhang SW, Chahl JS, Barth E, Venkatesh S (2000b) How honeybees make grazing landings on flat surfaces. Biol Cybern 83: 171–183

Stange G (1981) The ocellar component of flight equilibrium control in dragonflies. J Comp Physiol A 141: 335–347

Straw AD, Lee S, Dickinson MH (2010) Visual control of altitude in flying *Drosophila*. Curr Biol 20: 1–7

Straw AD, Rainsford T, O'Carroll DC (2008) Contrast sensitivity of insect motion detectors to natural images. J Vision 8: 1–9

Stuerzl W, Srinivasan MV (2010) Omnidirectional system with constant elevational gain and single viewpoint. Tenth Workshop on Omnidirectional Vision, Camera Networks and Sensors Zaragoza, Spain

Tammero LF, Dickinson MH (2002) The influence of visual landscape on the free flight behavior of the fruit fly *Drosphila melanogaster.* J Exp Biol 205: 327–343

Taylor GJ, Luu T, Ball D, Srinivasan MV (2011) Keeping up the pace: Honeybee flight speed regulation in a tethered flight arena. Proc, Australasian Soc for the Study of Animal Behaviour. ASSAB, Flinders University, Adelaide, p 57

Thurrowgood S, Moore RJD, Bland D, Soccol D, Srinivasan MV (2010) UAV attitude control using the visual horizon. Twelfth Australasian Conf on Robotics and Automation, Brisbane

Thurrowgood S, Soccol D, Moore RJD, Bland D, Srinivasan MV (2009) A vision based system for attitude estimation of UAVs. IEEE /RSJ Internat Conf on Intelligent Robots and Systems, ST. Louis, Missouri, USA

Todorovic S, Nechbya MC (2004) A vision system for intelligent mission profiles of micro air vehicles. IEEE Transactions on Vehicular Technology 53: 1713–1725

van Kleef J, James AC, Stange G (2005) A spatiotemporal white noise analysis of photoreceptor responses to UV and green light in the dragonfly median ocellus. J Gen Physiol 126: 481–497

Walker MM, Bitterman ME (1985) Conditioned responding to magnetic fields by honey bees. J Comp Physiol A 157: 67–73

Weber K, Venkatesh S, Srinivasan MV (1997) Insect inspired behaviours for the autonomous control of mobile robots. From living eyes to seeing machines. Oxford University Press, U. K., pp 226–248

Wehner R, Labhart T (2006) Polarization vision In: Warrant E, Nilsson D-E (eds) Invertebrate vision. Cambridge University Press, Cambridge, U. K, pp 291–348

Wittlinger M, Wehner R, Wolf H (2007) The desert ant odometer: a stride integrator that accounts for stride length and walking speed. J Exp Biol 210: 198–207

Yagi Y, Nishizaw Y, Yachida M (1995) Map-based navigation for a mobile robot with omnidirectional image sensor COPIS. IEEE Transactions on Robotics and Automation 11: 634–648

Zeil J, Nalbach G, Nalbach H-O (1986) Eyes, eyestalks and the visual world of semi-terrestrial crabs. J Comp Physiol A 159: 801–811

Vision

II

Color sensors of butterflies

3

Kentaro Arikawa

Contents

Abstract

Colorful butterflies are strongly visual animals, and have sophisticated color vision. The compound eyes of the Japanese yellow swallowtail butterfly, *Papilio xuthus*, contain ultraviolet, violet, blue, green, red and a broad-band class of receptors, which are embedded in the ommatidia in three fixed combinations. The eyes are, therefore, a mesh of three types of spectrally-heterogeneous ommatidia. Given that the eyes of *Papilio* are equipped with six classes of spectral receptors, their color vision may be hexachromatic. The foraging *Papilio* can only discriminate a one nanometer difference at three wavelength regions around 430, 480 and 560 nm, indicating that their color vision is actually tetrachromatic. The noise-limited color opponency model has predicted that the tetrachromacy is based on the ultraviolet, blue, green and red receptors. These receptors are contained together in at least one type of ommatidia, where they form a single rhabdom. In principle, such an organization makes the butterfly able to discriminate colors at the single-pixel level. In fact, a foraging butterfly can discriminate between a colored disk and a gray one at their visual subtense angle of around 0.7–1.0 degree, which is close to the limit of the spatial resolution predicted for their eye optics.

Kentaro Arikawa
Graduate University for Advanced Studies (Sokendai), Laboratory of Neuroethology
Shonan Village
Hayama, Kanagawa 240-0193, Japan
e-mail: arikawa@soken.ac.jp

1. Introduction: Basic design of insect eyes

The visual systems of insects are equipped with compound eyes, which are composed of thousands of optical units, the ommatidia, each gathering light from a small visual field. Larger eyes contain more ommatidia than smaller eyes, as, for example, the eye of the large Japanese Yellow Swallowtail butterfly, *Papilio xuthus*, has more than 12 000 ommatidia, while an eye of the Small White, *Pieris rapae*, has about 6 000. Due to their structure, compound eyes are generally divided into two classes, the apposition and super-

position eyes, characteristic of diurnal and nocturnal species, respectively. This chapter will focus on the appositional eyes of butterflies (for a detailed description of superposition eyes, see Nilsson 1989).

The ommatidium of the apposition eye of butterflies has distally an optical system which projects incident light onto a bundle of photoreceptors. The cell membrane of each photoreceptor has microvillar projections located near the center of the ommatidia, and together are called the rhabdomere. Within each ommatidium, the rhabdomeres form a long, fused rhabdom. Figure 1a presents

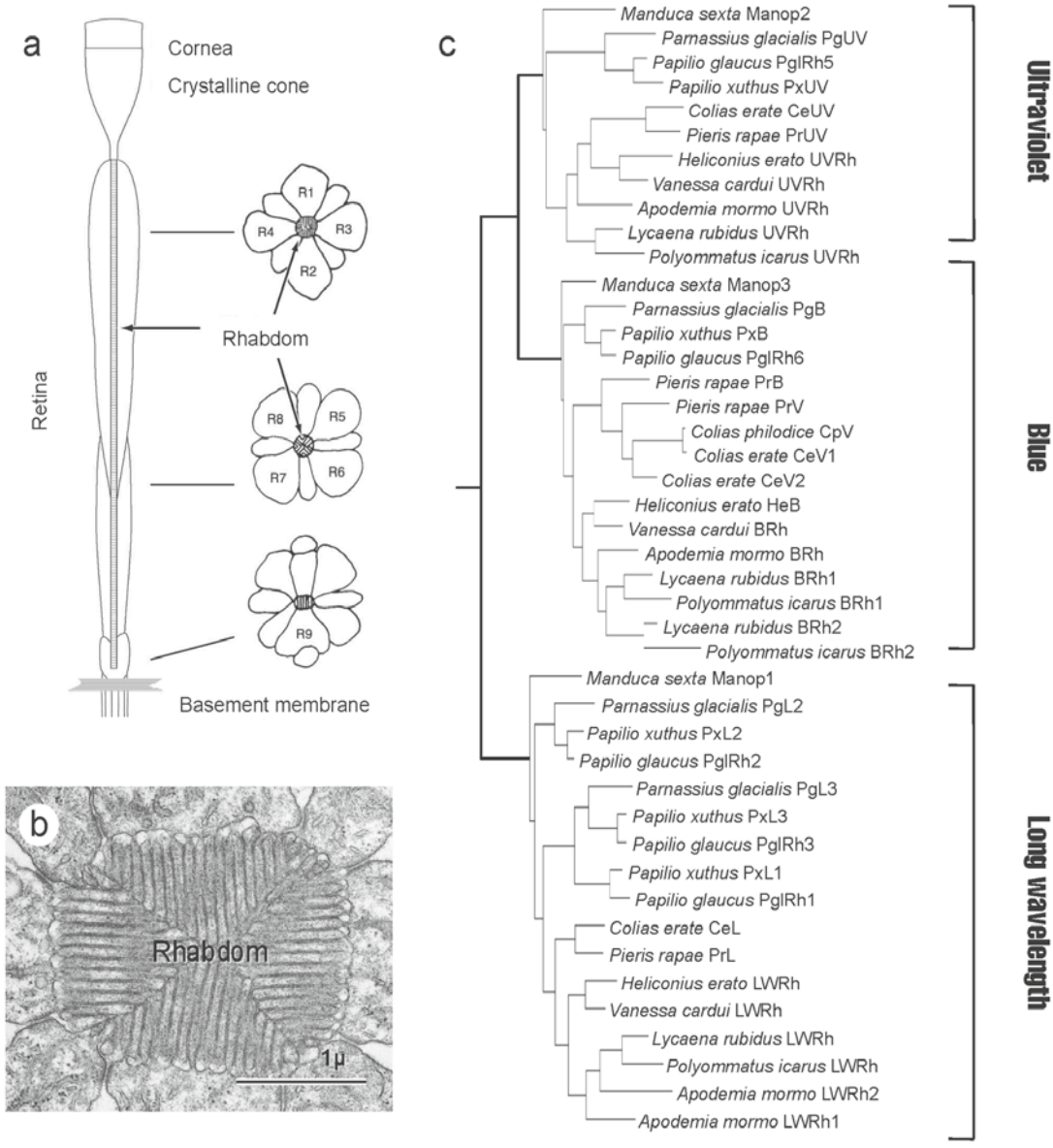


Fig. 1 **a** Schematic diagram of a *Papilio* ommatidium. R1-R9; photoreceptors 1–9. **b** Electron micrograph of a transverse section of the rhabdom through the distal tier. **c** Phylogenetic relationship of insect opsins based on the amino acid sequences deduced from the full length cDNA sequences

a diagram of an ommatidium of *Papilio xuthus*, which contains nine photoreceptor cells, R1 to R9. In *Papilio*, as well as in many other butterflies, the distal two-thirds of the rhabdom is composed of the rhabdomeres of R1–R4 (Fig. 1b). The proximal one-third of the rhabdom consists of the rhabdomeres of R5–R8. The ninth photoreceptor, R9, bears a few microvilli at the base of the rhabdom. At the basement membrane, the photoreceptor cells taper and form axons projecting to the lamina, where they form synaptic connections with second order interneurons.

The microvilli are tubular extensions of the photoreceptor cell membranes (Fig. 1b). The cell membranes are lipid bilayers, and consequently the rhabdoms have a larger refractive index than the surrounding cytoplasm, and the rhabdom therefore acts as an optical waveguide. Light entering at the distal tip of the rhabdom propagates along the rhabdom, even if the rhabdom is physically not straight; the ommatidia are, therefore, optically-independent units, each sampling a small fraction of the visual environment.

The microvillar membranes contain abundant rhodopsins, the visual pigments that trigger the phototransduction process. While propagating along the rhabdom, the light is absorbed by the rhodopsins. A rhodopsin molecule is composed of an opsin protein and a chromophore, the 11-*cis* form of a vitamin A aldehyde, retinal. Upon light absorption, the 11-*cis* retinal is transformed to the all-*trans* form, which then triggers the conversion of the rhodopsin to its active form, metarhodopsin. The latter visual pigment state then activates the phototransduction cascade, the end result of which is the receptor potential, a depolarization of the photoreceptor's membrane potential.

The absorption spectrum of a rhodopsin crucially depends on the animo acid sequence of opsin, which determines the photochemical interaction of the chromophore and the opsin protein. For example, honeybee eyes express three independent opsin genes, encoding opsins of an ultraviolet- (UV-), blue- (B-) and long wavelength- (L-) rhodopsin, respectively (Fig. 1c). These rhodopsins all have the same chromophore, 11-*cis* retinal, so that the difference in the spectral properties is due to the difference in amino acid sequences. More specifically, a few amino acid residues close to the chromophore and acting as counter ions, play a crucial role in determining the spectral property of the rhodopsin (Terakita et al. 2004; Wakakuwa et al. 2010).

Accumulating evidence indicates that insect opsins can be divided into three categories, belonging to UV-, B- and L-absorbing types (Fig. 1c). Probably the basic design of compound eyes is to have one opsin from each category, as in honeybees, but variation on this basic design is quite common. For example, the Small White, *Pieris rapae*, has two B-opsins in addition to one UV-opsin and one L-opsin (Wakakuwa et al. 2004; Arikawa et al. 2005). Two Swallowtail species of the genus *Papilio* have at least three L-opsins in the eye, while they have single UV- and B-opsins (Briscoe 2000; Arikawa 2003). On the other hand, the red flour beetle *Tribolium castaneum* has only two types of opsins; it lacks the B-opsin (Jackowska et al. 2007).

2. The eye of *Papilio*

2.1 Spectral organization of the eye

2.1.1 Photoreceptor spectral sensitivity

Insects that have rhodopsins from each of the three categories, UV, B and L, may have the capacity for trichromatic color vision. In fact, honeybees have a trichromatic system

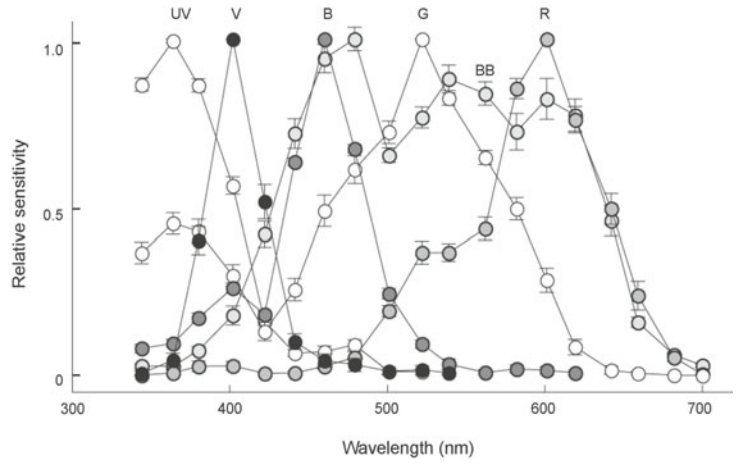


Fig. 2 Spectral sensitivities of single photoreceptors in the *Papilio* eye. UV, ultraviolet; V, violet; B, blue; G, green; R, red; BB, broad-band

based on the UV, B and G receptor set in the eye. Nevertheless, we now know that there are many species that have more than three classes of spectral receptors. *Papilio xuthus* is the earliest case where this has been firmly demonstrated.

The spectral sensitivity of a given photoreceptor can be determined by intracellular recording of the receptor potential in response to a series of monochromatic lights. We thus identified at least six classes of spectral receptors in the eye of *Papilio xuthus*: a UV, violet (V), B, G, red (R) and broad-band (BB) class (Fig. 2). These receptors are all embedded in the ommatidia, occupying a unique position there: we found that the photoreceptors R1 and R2 are either UV, V or B receptors, R3 and R4 are G receptors and R5–8 are either G, R or BB receptors. The R9 photoreceptors are probably either G or R (Arikawa 2003).

The results of combined electrophysiology, optophysiology, anatomy and molecular biology have demonstrated that the ommatidia are spectrally heterogeneous. In *Papilio*, the ommatidia are histologically heterogeneous as follows from the characteristic pigmentation around the rhabdom. The color of the pigment is red in about two-thirds of ommatidia, whereas the remaining one-third has a yellow pigment (Fig. 3a). Although these pigments exist outside the rhabdom, they act as spectral filters by absorbing the boundary wave of the light propagating along the rhabdom. The filtering effect is observable when the eye is illuminated from inside: the ommatidia look red or yellow (Fig. 3b). Some red ones have a less saturated pinkish appearance. These ommatidia have another conspicuous feature, that is, they fluoresce under UV excitation light (Fig. 3c). The fluorescing material is probably 3-hydroxyretinol, concentrated in the rhabdom near its distal tip. Taken together, the *Papilio* eye contains at least three types of ommatidia, namely red-pigmented, red/fluorescing-pigmented and yellow-pigmented (Arikawa 2003).

2.1.2 Opsin expression pattern

To find out which classes of spectral receptors are contained in the ommatidia of specific pigmentation, we adopted a molecular biological approach. From poly-A RNA extracted from the *Papilio* eye, we eventually cloned five full-length cDNAs encoding opsins. Phylogenetic analysis has indicated that the opsins are: one UV-opsin (PxUV), one B-opsin (PxB) and three L-opsins (PxL1–3, Fig. 1c). We localized the mRNAs in

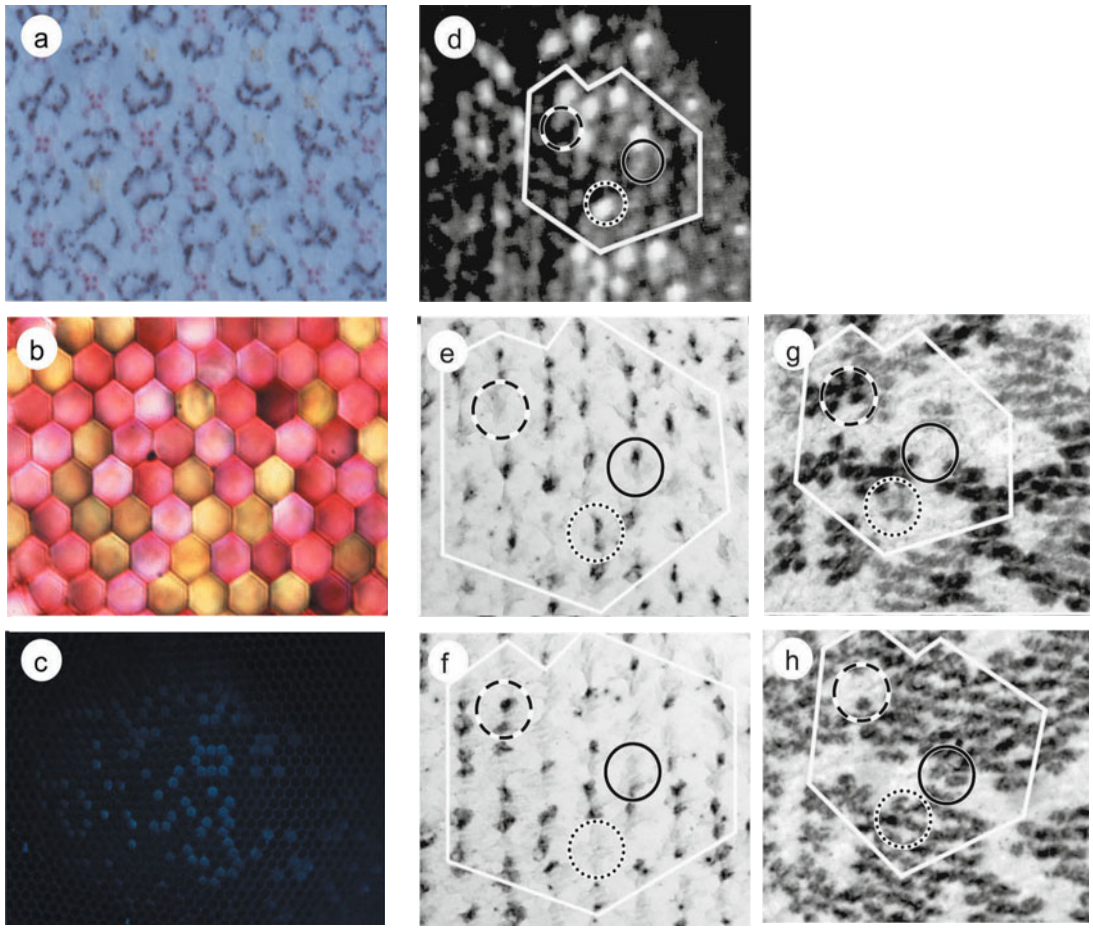


Fig. 3 Anatomy of the *Papilio* eye. **a** Unstained transverse section showing four clusters of red or yellow pigment surrounding the rhabdom in each ommatidium. **b** An eye illuminated from inside with white light. The red or yellow color of the ommatidia corresponds to the pigment color around the rhabdom. **c** UV-induced fluorescence demonstrating a subset of ommatidia emitting a whitish fluorescence. **d** Fluorescence picture of a frozen section under UV excitation. **e** Localization of PxUV mRNA by *in situ* hybridization. **f** Localization of PxB. **g** Localization of PxL2 (G-opsin). **h** Localization of PxL3 (R-opsin). The white polygons in **d–h** indicate the same set of ommatidia. Solid circle, type I ommatidium; dotted circle, type II ommatidium; broken circle, type III ommatidium. See Table 1

the eye by *in situ* hybridization, the results of which are shown in Fig. 3d–h. These five pictures were taken from consecutive frozen sections collected from one eye. Immediately after the sections were cut, we took a fluorescence picture of a section under UV excitation to identify the fluorescing ommatidia: the whitish ones are the fluorescing ommatidia (Fig. 3d). Other sections were then labeled with probes specifically hybridizing the PxUV (Fig. 3e), PxB (Fig. 3f), PxL2 (Fig. 3g) and PxL3 (Fig. 3h) mRNAs, respectively.

The expression pattern of PxUV and PxB is straightforward: they are expressed in either R1 or R2, and the pattern is complementary in all three possible combinations: UV/B, UV/UV and B/B. The PxL2 and PxL3 mRNAs are found in R5–8, but the labeling pattern is rather complicated. In the PxL2 labeling, there are ommatidia whose R5–8 are densely labeled, weakly labeled and not

labeled. However the PxL3 labeling has only two patterns; R5-R8 are all labeled or not at all labeled. In the serial sections we correlated the labeling pattern. The distal UV/B type has the proximal R5–8 expressing PxL3. The B/B type has R5–8 expressing PxL2. The UV/UV type fluoresces and has R5–8 expressing both PxL2 and PxL3 (Fig. 3).

Although not shown here, PxL2 was found also in the R3 and R4 in all ommatidia. The R3 and R4 are G receptors, so PxL2 must be a G-opsin. Interestingly, the PxL2-expressing R3 and R4 co-express PxL1 in the ventral region of the eye: PxL1 is found only in these locations. This is rather enigmatic because if the PxL1 is functional in these cells, it is also a G-opsin whose spectral property must be almost identical to PxL2. The question of the functional significance of PxL1 still remains to be addressed.

2.1.3 Double expression and filter function

Particularly interesting is the property of the fluorescing ommatidia, which express UV-opsin in R1 and R2, and co-express G-opsin (PxL2) and R-opsin (PxL3) in R5–R8. These receptors were identified as V (R1,2) and BB receptors (R5–8), respectively (Fig. 2).

The spectral sensitivity of the V receptors has an aberrantly narrow profile with peak wavelength 400 nm (Fig. 2). However, they express the UV-opsin present in the UV receptors peaking at 360 nm (Fig. 3e). The modified shape of the V receptor's spectral sensitivity is due to the fluorescing material acting as a UV absorbing filter. We calculated the filtering effect by assuming that the fluorescing material maximally absorbs 330 nm light and has enough optical density to suppress the incident light in the UV. The reduction of UV content results in a lowered responsiveness of the receptors to UV, so that the spectral sensitivity is shifted towards longer wavelengths (Arikawa et al. 1999).

The mechanism underlying the BB receptors is even more interesting. Their reduced sensitivity in the UV is also attributed to the effect of the fluorescing filter. The expanded sensitivity in the visible light range is explained by postulating that both G- and R-opsins are functional. Actually a long-time accepted "dogma" in vision research was that one photoreceptor contains one type of rhodopsin. The *Papilio* BB receptors are the first example of photoreceptors violating the dogma (Arikawa et al. 2003). Yet, we currently only know that the *Papilio* BB receptors express the mRNAs of G- and R-opsins, but whether the two rhodopsins are indeed functional remains to be confirmed. Photoreceptors expressing two or more opsins, or at least containing more than one opsin mRNA, were also found in other animals (Sakamoto et al. 1996; Sison-Mangus et al. 2006; Jackowska et al. 2007; Mazzoni et al. 2008; Stavenga and Arikawa 2008; Awata et al. 2009), but solid electrophysiological proof of an expanded spectral sensitivity is lacking in all cases.

2.1.4 Spectral heterogeneity of ommatidia

Ommatidial heterogeneity has already been noticed for decades in butterflies, wasps and flies (Bernard and Miller 1970; Ribi 1978; Franceschini et al. 1981), but its careful analysis has been completed for the first time in *Papilio xuthus*. There are at least three types of ommatidia containing different sets of spectral receptors (Table 1), and, at least locally, they distribute randomly (Arikawa and Stavenga 1997). Ommatidial heterogeneity and local randomness are probably widespread (Wakakuwa et al. 2006; Briscoe 2008), but the spectral sensitivities of photoreceptors and opsin expression pattern seem strongly species-dependent (Wakakuwa et al. 2006).

Table 1 Three types of ommatidia of the *Papilio* compound eye. UV, ultraviolet; V, violet; B, blue; DG, double-peaked green with secondary peak in the UV (see Fig. 3); SG, single-peaked green without the UV peak (Arikawa et al. 1999); R, red; BB, broadband

Type of ommatidia:	**Type I**		**Type II**		**Type I**	
Pigmentation:	**Red**		**Red**		**Yellow**	
Fluorescence:	**No**		**Yes**		**No**	
Photoreceptors:	***S*(λ)**	***opsin***	***S*(λ)**	***opsin***	***S*(λ)**	***opsin***
R1	UV	PxUV	V	PxUV	B	PxB
R2	B	PxB	V	PxUV	B	PxB
R3–4	DG	PxL1+L2	SG	PxL1+L2	DG	PxL1+L2
R5–8	R	PxL3	BB	PxL2+L3	DG	PxL2
R9	R?	?	R?	?	DG?	?

3. Color vision of *Papilio*

3.1 Color discrimination and its spatial size limit

3.1.1 Color vision

The existence of a variety of spectral receptors in the compound eye implies that *Papilio* butterflies see colors, but this requires behavioral proof. Karl von Frisch was the first to convincingly demonstrated color vision in an insect, using a series of color plates (Frisch 1914). Basically we applied his protocol for honeybees to the feeding behavior of *Papilio*. Firstly we trained a *Papilio* to take sucrose solution placed on a paper disk of a certain color. When the butterfly became adept at visiting the disk spontaneously, we presented four disks of different colors including the training color. The butterfly shown in Fig. 4 was trained towards blue, and correctly visited the blue disk among four different colors. The butterfly also chose the blue disk when presented with seven disks of different densities of gray. If the butterfly visual system is monochromatic, the butterfly would have confused the blue disk and one of the gray disks that have similar subjective brightness. But this did not happen, which indicates that the butterfly used color vision when searching the food source (Kinoshita et al. 1999)

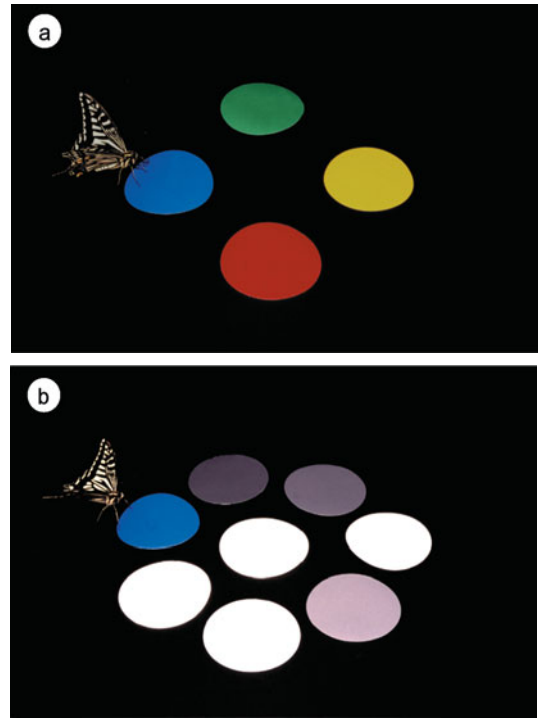


Fig. 4 Demonstration of color vision in foraging *Papilio*

3.1.2 Experimental procedure using an Y-maze

Humans have R-, G- and B-cone photoreceptors, which are densely packed in the central part of the retina. If these cones are stimulated simultaneously and their signals properly processed, the sense of color will be produced. In reverse, if the angular size of a target is as small as the angle subtended by a single cone, i. e. at around the limit of spatial resolution, the sense of color will be unreliable (Brainard et al. 2008). In apposition compound eyes, a single ommatidium with a fused rhabdom provides a single pixel of the image obtained by the eye. The fundamental difference between a butterfly ommatidium and a vertebrate cone is that the ommatidium contains multiple classes of spectral receptors. Therefore it is theoretically possible for *Papilio* to detect the color of a target whose size is close to the eye's spatial resolution limit.

How small can the visual target be for *Papilio* to detect its color? We asked this question using a series of dual-choice experiments and adopting a Y-maze apparatus (Fig. 5a). The stimuli 1 and 2 of the Y-maze were two disks of the same diameter, one in the training color and another in gray. We trained a *Papilio* in the training area to take sucrose solution in front of a colored disk presented vertically on the removable wall separating the selection and the training areas. We then released the trained butterfly in the training area, removed the wall and let them fly into the selection area. The butterfly could see both stimuli 1 and 2 as long as it was in the region shaded in Fig. 5a. When the butterfly could distinguish the colored from the gray stimulus, the butterfly went into the arm, and when the butterfly crossed the border between the selection area and the arms, we counted the event as either a positive or a negative response. After entering the correct arm the butterflies approached the target and touched it with the proboscis. We systematically changed the diameter of the disks to determine the minimum visual angle for discriminating the colors of the disks (Takeuchi et al. 2006).

3.1.3 Size limit of color discrimination

Two 5 cm diameter disks presented at a distance of 50 cm from the border between the selection area and the arm, thus having an angular subtense of 5.7°, were correctly discriminated between (Fig. 5b). However, as the disk size decreased down to 1°, the correct choice ratio became around 50 %. Assuming that the butterflies positively discerned the color of the disk when the correct choice ratio was larger than 60 %, it appeared that *Papilio* could correctly select colored disks when the subtense angle was larger than about 1° (Fig. 5b).

A visual angle of 1° roughly corresponds to the spatial resolving power of *Papilio* predicted from the anatomy of the eye (Fig. 5c). The highest sampling frequency that a compound eye can resolve, v_s, is that for which there is one receptor unit for each half cycle of the grating (Land 1997). It is given by $v_s = 1/(2\Delta\phi)$, where $\Delta\phi$ is the interommatidial angle. The interommatidial angle of the *Papilio* eye is about 1.0° (Arikawa, unpublished observation), so that v_s is about 0.5 cycle per degree. This means that *Papilio* can theoretically resolve two 1° dots separated by a spatial interval of 1°. However, the effective cut-off frequency of the optics, v_{opt}, is affected by the ommatidial acceptance angle, $\Delta\rho$. The electrophysiologically determined $\Delta\rho$ of a *Papilio* ommatidium is about 1.9° (Arikawa, unpublished; see also (Horridge et al. 1983) for *Papilio aegeus*). Note that $\Delta\rho$ is about twice the value of $\Delta\phi$ (1.0°). The relationship $\Delta\rho = 2\Delta\phi$ follows from $v_s = v_{opt}$ (Land 1997). The spatial resolution of *Papilio* is thus estimated to be about 1.0°, and the butterfly can therefore probably detect the

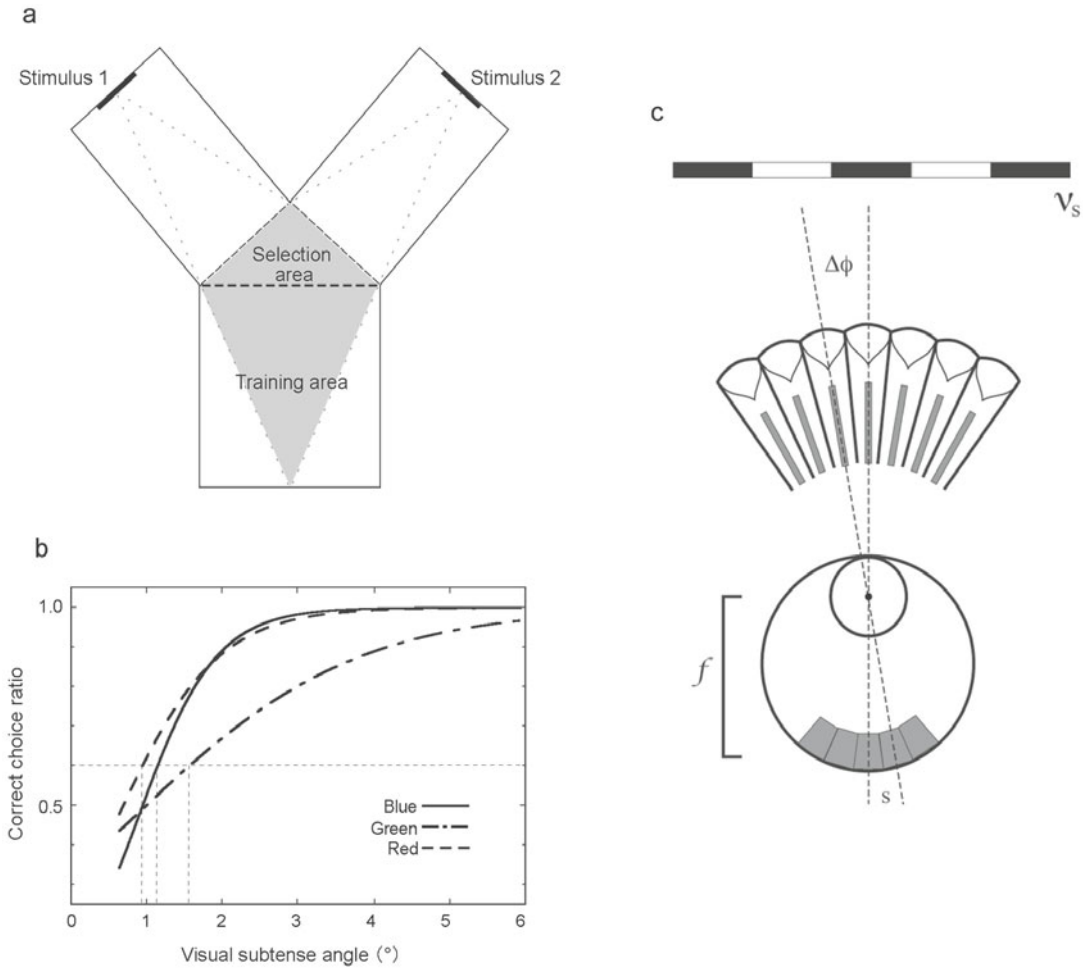


Fig. 5 **a** Y-maze apparatus. The tested butterflies were able to see two targets simultaneously while they were in the shaded area. **b** Relationship of the correct choice ratio and the angular subtense of the stimuli. **c** Relationship between the visual acuity in a camera eye (below) and an apposition compound eye (above). *s*, receptor separation, *f*, focal length; $\Delta\phi$, interommatidial angle; ν_s, spatial frequency of stimulus (modified from Land 1997)

color of a target whose size is close to the spatial resolution limit (Takeuchi et al. 2006).

3.2 Wavelength discrimination

3.2.1 Behavioral measure of wavelength discrimination in *Papilio*

Given that *Papilio* butterflies have six classes of spectral receptors in the retina, is their color vision hexachromatic? This can be answered at least in part by measuring the wavelength discrimination ability, that is, the threshold of detecting the difference in wavelength of two monochromatic lights as different colors. The minimum wavelength difference that humans can detect, $\Delta\lambda$, is extreme, ~1 nm, in the wavelength regions around 500 and 600 nm. Because the human color vision system is trichromatic, the $\Delta\lambda$ function has two regions with high color discrimination, or "troughs".

To predict the number of spectral recep-

tor classes involved in *Papilio* color vision, we measured the $\Delta\lambda$ function of *Papilio* using their proboscis extension behavior. We then applied a series of monochromatic lights. A *Papilio* was first trained to take sucrose solution at a light target of a certain wavelength. The trained butterfly was then forced to select one of two monochromatic lights. The wavelength of one light was always the training wavelength, λ_{tr}, and another light, the test light λ_{te}, was changed in the range $\lambda_{tr} - \Delta\lambda < \lambda_{te} < \lambda_{tr} + \Delta\lambda$, with $0 \leq \Delta\lambda \leq 20$ nm. Figure 6a shows the results obtained from 480 nm-trained butterflies (λ_{tr} = 480 nm). When λ_{te} was 480 nm, the butterflies faced two identical lights, and therefore their selection was random, i. e. 50 %. The choice ratio increased to 90 % when the wavelength difference increased, making the choice curve V-shaped. We estimated the $\Delta\lambda$ by assuming that the butterflies could discriminate between the two lights when the correct choice ratio was 60 %. Figure 6b is a plot of the $\Delta\lambda$ value for 16 training wavelengths. Clearly there are three troughs, indicating a tetrachromatic system (Koshitaka et al. 2008).

3.2.2 Model calculation

Which four classes among the six classes of receptors present contribute here? We modified the noise-limited color opponency model of Vorobyev and Osorio (1998), and applied it to predict the wavelength discrimination property. The modified model is based on the assumption that visual thresholds are set by noise originating in the photoreceptors and that intensity-based cues are not used for color discrimination. When the perceptual distance between stimuli predicted by the model is less than a threshold distance, the stimuli of different wavelength cannot be discriminated. For details, see Koshitaka et al. (2008).

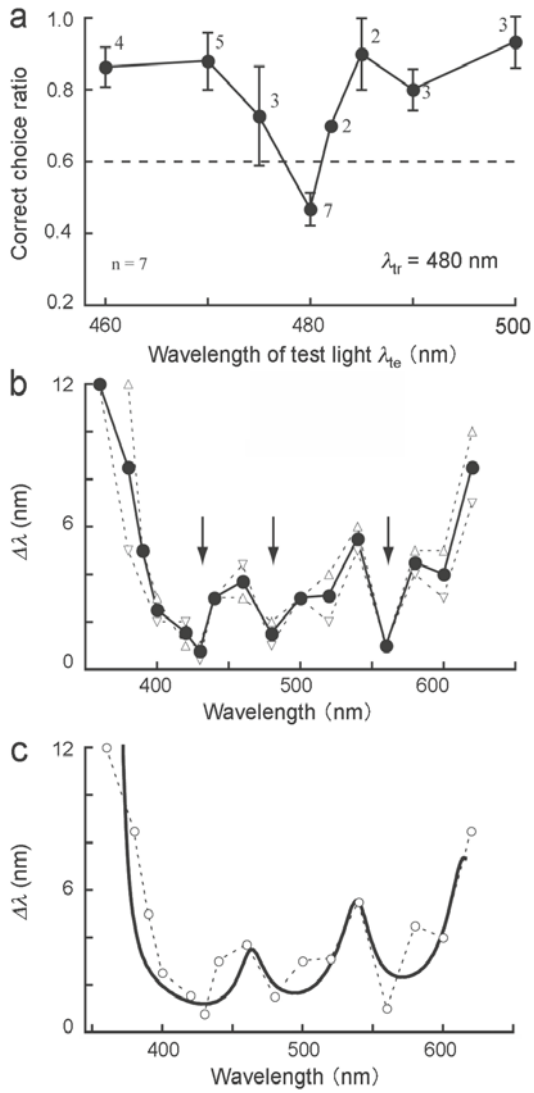


Fig. 6 **a** Wavelength discrimination of 480 nm-trained butterflies. The numbers at the data points indicate the number of butterflies tested. **b** The $\Delta\lambda$function. The curve exhibits three troughs (arrows) at 430, 480 and 560 nm, respectively. **c** Model calculation with UV, B, G and R receptors (solid line) and the behavioural data (dotted lines, see Fig. 9b)

When incorporating all six receptor classes, UV, V, B, G, R and BB, the model-predicted $\Delta\lambda$ function had four troughs, in conflict with the behavioral data. Removal of the BB receptor did not improve the match. Further removal of the UV receptor

eliminated the first trough in the UV region, but the match was still not optimal. The best match was achieved by removing the V, BB and SG (single-peaked G receptors specific to type II ommatidia, see Table 1) receptors (Fig. 6c). We therefore concluded that the color vision of foraging *Papilio* is tetrachromatic, based on the set of UV, B, G and R receptors (Koshitaka et al. 2008).

Of particular interest is the fact that the two excluded receptors, V and BB, are located in the fluorescing ommatidia (see Table 1). Fluorescing ommatidia are possibly dedicated for some functions other than color vision.

Concluding remarks

This chapter has focused on the visual system of *Papilio xuthus* specifically concerning its color vision characteristics. The eyes of *Papilio* are extremely complex, with six classes of receptors embedded in three different types of ommatidia that are rather randomly distributed. At the molecular level, the *Papilio* eye expresses five opsins, some of which are unexpectedly found to be co-expressed. The immediate question arising now is what are the features shared by other species and what are the features specific to *Papilio*. In fact, the honeybee eye, with three classes of receptors and three corresponding opsins, seems somewhat simpler by comparison (Wakakuwa et al. 2005). To extract general features across species, we have to compare many species. Such a comparative approach will significantly enhance our understanding of evolution, which remains one of the most important topics in modern biology. Having this scope in mind, we have embarked upon a program of comparative studies on related butterflies (Stavenga and Arikawa 2006).

As the starting point of the comparative work, we investigated the Small White, *Pieris rapae*, a pierid. Its eyes are equipped with six classes of spectral receptors, which are embedded in the ommatidia in three fixed combinations, as in *Papilio*. However, prominent differences have become evident: the spectral sensitivity profiles of the photoreceptors are quite different. The most conspicuous feature of the *Pieris* eye is its sexual dimorphism. In female eyes, the photoreceptors are of the UV, V, B, G, R and dark-red (DR) class, but the male eyes lack V receptors and instead have double-peaked B (dB) receptors. At the molecular level, the female V receptors and male dB receptors share the PrV opsin, a *Pieris rapae* opsin clustered in the B opsin category (Fig. 2). The mechanism underlying this sexual dimorphism is attributed to the filtering effect of a male-specific fluorescing pigment, which fluoresces under 420 nm light. By absorbing 420 nm light, the pigment reduces the sensitivity of the V receptors at 420 nm, making the spectral sensitivity double-peaked.

Another notable feature of the *Pieris* eye is that it has another B-opsin, PrB, in addition to PrV, which is due to the duplication of the B-opsin gene. The subfunctionalization of the duplicate occurred in the lineage of *Pieris*. The family Pieridae consists of four subfamilies, Pierinae, Coliadinae, Dismorphinae and Pseudopontiinae. In addition we have analyzed the eyes of the Pale Clouded Yellow, *Colias erate*, a species of the Coliadinae subfamily, and found that they also have at least two B-opsins, CeV1 and CeV2. These two B-opsins of *Colias* clustered with PrV, not with PrB. The phylogenetic relationship of these opsins suggests that duplication of the ancestral B-opsin occurred before the Pierinae and Coliadinae diverged. Then, in the lineage of Coliadinae, the B-opsin was disfunctionalized and the V-opsin again duplicated (Awata et al. 2009; Wakakuwa et al. 2010). We are now in the process of the molecular characterization of the opsins of several other subfamilies, and the emerging impression is that for the opsins, the subfamily is the important taxonomic unit.

Despite the pronouced species specificity, the feature shared by all investigated species is that each ommatidium contains at least two different spectral receptors. In principle, this makes the wavelength analysis possible at the single-pixel level, which is functionally similar to the recently-developed hyper-spectral imaging cameras. However, even the most recent model of such camera includes a scanning mechanism that moves the internal grating. Because of this mechanism, it usually takes 10–20 sec to acquire a set of data, and, therefore, these cameras cannot be applied to any moving objects. This problem can be solved if we develop a wavelength discriminating device of

one micrometer-order in diameter, and densely pack millions of them together to make a single multi-spectral sensor. The butterfly rhabdom, whose diameter is 1–2 µm, would provide an ideal model for such a sensor.

References

Arikawa K (2003) Spectral organization of the eye of a butterfly, *Papilio*. J Comp Physiol A 189: 791–800

Arikawa K, Mizuno S, Kinoshita M, Stavenga DG (2003) Coexpression of two visual pigments in a photoreceptor causes an abnormally broad spectral sensitivity in the eye of a butterfly, *Papilio xuthus*. J Neurosci 23: 4527–4532

Arikawa K, Mizuno S, Scholten DG, Kinoshita M, Seki T, Kitamoto J, Stavenga DG (1999) An ultraviolet absorbing pigment causes a narrow-band violet receptor and a single-peaked green receptor in the eye of the butterfly *Papilio*. Vision Res 39: 1–8

Arikawa K, Stavenga DG (1997) Random array of colour filters in the eyes of butterflies. J Exp Biol 200: 2501–2506

Arikawa K, Wakakuwa M, Qiu X, Kurasawa M, Stavenga DG (2005) Sexual dimorphism of short-wavelength photoreceptors in the small white butterfly, *Pieris rapae crucivora*. J Neurosci 25: 5935–5942

Awata H, Wakakuwa M, Arikawa K (2009) Evolution of color vision in pierid butterflies: blue opsin duplication, ommatidial heterogeneity and eye regionalization in *Colias erate*. J Comp Physiol A 195: 401–408

Bernard GD, Miller WH (1970) What does antenna engineering have to do with insect eyes? IEEE Student J 8: 2–8

Brainard DH, Williams DR, Hofer H (2008) Trichromatic reconstruction from the interleaved cone mosaic: Bayesian model and the color appearance of small spots. J Vis 8: 15 11–23

Briscoe AD (2000) Six opsins from the butterfly *Papilio glaucus*: Molecular phylogenetic evidence for paralogous origins of red-sensitive visual pigments in insects. J Mol Evol 51: 110–121

Briscoe AD (2008) Reconstructing the ancestral butterfly eye: focus on the opsins. J Exp Biol 211: 1805–1813

Franceschini N, Kirschfeld K, Minke B (1981) Fluorescence of photoreceptor cells observed *in vivo*. Science 213: 1264–1267

Frisch K von (1914) Der Farbensinn und Formensinn der Biene. Zool J Physiol 37: 1–238

Horridge GA, Marcelja L, Jahnke R, Matic T (1983) Single electrode studies on the retina of the butterfly *Papilio*. J Comp Physiol A 150: 271–294

Jackowska M, Bao R, Liu Z, McDonald EC, Cook TA, Friedrich M (2007) Genomic and gene regulatory signatures of cryptozoic adaptation: Loss of blue sensitive photoreceptors through expansion of long wavelength-opsin expression in the red flour beetle *Tribolium castaneum*. Front Zool 4: 24

Kinoshita M, Shimada N, Arikawa K (1999) Colour vision of the foraging swallowtail butterfly *Papilio xuthus*. J Exp Biol 202 (Pt 2): 95–102

Koshitaka H, Kinoshita M, Vorobyev M, Arikawa K (2008) Tetrachromacy in a butterfly that has eight varieties of spectral receptors. Proc Biol Sci 275: 947–954

Land MF (1997) Visual acuity in insects. Annu Rev Entomol 42: 147–177

Mazzoni EO, Celik A, Wernet MF, Vasiliauskas D, Johnston RJ, Cook TA, Pichaud F, Desplan C (2008) Iroquois complex genes induce co-expression of rhodopsins in *Drosophila*. PLoS Biol 6: e97

Nilsson D-E (1989) Optics and evolution of the compound eye. In: Stavenga DG, Hardie RC (eds) Facets of vision. Springer-Verlag, Berlin Heidelberg New York London Paris Tokyo, pp 30–73

Ribi WA (1978) A unique hymenopteran compound eye. The retina fine structure of the digger wasp *Sphex cognatus* Smith (Hymenoptera, Sphecidae). Zool Jb Anat Bd 100: 299–342

Sakamoto K, Hisatomi O, Tokunaga F, Eguchi E (1996) Two opsins from the compound eye of the crab *Hemigrapsus sanguineus*. J Exp Biol 199: 441–450

Sison-Mangus MP, Bernard GD, Lampel J, Briscoe AD (2006) Beauty in the eye of the beholder: the two blue opsins of lycaenid butterflies and the opsin gene-driven evolution of sexually dimorphic eyes. J Exp Biol 209: 3079–3090

Stavenga DG, Arikawa K (2006) Evolution of color and vision of butterflies. Arthropod Struct Dev 35: 307–318

Stavenga DG, Arikawa K (2008) One rhodopsin per photoreceptor: *Iro-C* genes break the rule. PLoS Biol 6: e115

Takeuchi Y, Arikawa K, Kinoshita M (2006) Color

discrimination at the spatial resolution limit in a swallowtail butterfly, *Papilio xuthus*. J Exp Biol 209: 2873–2879

Terakita A, M. K, Tsukamoto H, Yamashita T, Miyata T, Shichida Y (2004) Counterion displacement in the molecular evolution of the rhodopsin family. Nature Struct Molec Biol 11: 284–289

Vorobyev M, Osorio D (1998) Receptor noise as a determinant of colour thresholds. Proc Biol Sci 265: 351–358

Wakakuwa M, Kurasawa M, Giurfa M, Arikawa K (2005) Spectral heterogeneity of honeybee ommatidia. Naturwissenschaften 92: 464–467

Wakakuwa M, Stavenga DG, Arikawa K (2006) Spectral organization of ommatidia in flower-visiting insects. Photochem Photobiol 83: 27–34

Wakakuwa M, Stavenga DG, Kurasawa M, Arikawa K (2004) A unique visual pigment expressed in green, red and deep-red receptors in the eye of the small white butterfly, *Pieris rapae crucivora*. J Exp Biol 207: 2803–2810

Wakakuwa M, Terakita A, Koyanagi M, Stavenga DG, Shichida Y, Arikawa K (2010) Evolution and mechanism of spectral tuning of blue-absorbing visual pigments in butterflies. PLoS ONE 5: e15015

Insect tangential cell analogues and implications for efficient visuomotor control

4

J. Sean Humbert, Andrew M. Hyslop

Contents

Abstract

In this chapter, a mathematical analogue for tangential cell processing is proposed and a framework is developed to demonstrate how the outputs of tangential cells from the horizontal (HS) system can be appropriately amplified and combined to generate motor commands to achieve reflexive navigation behavior. It is shown that the resulting feedback synthesis task can be cast as a combined static state estimation and linear feedback control problem.

J. Sean Humbert
University of Maryland
Department of Aerospace Engineering
3181 Glenn L Martin Hall
College Park, MD 20742, USA
e-mail: humbert@umd.edu

1. Introduction

Over the past 5 years there has been significant research investment and progress in actuation and fabrication technology for micro-scale systems (Wood et al 2008). Sensors, processing, and feedback control architectures, on the other hand, are dramatically behind the curve at these scales. This is particularly true for flying microrobotic systems, where the challenges due to the high bandwidth dynamics and stringent size, weight, and power constraints render traditional technology and paradigms unusable. Novel sensors and sensory processing architectures that enable stable flight control, obstacle avoidance, and detection of external stimuli will need to be developed if autonomous microsystems are to ultimately be realized.

A promising research direction is biologically inspired approaches (Barrows et al 2003) for navigation on micro-air vehicles as an alternative to traditional paradigms. Of particular interest is *optic flow*, the visual motion produced by body rotation and translation. Patterns of optic flow encode rigid body motions, and are also the primary source of relative distance to surrounding objects through translation-induced parallax (Gibson 1950). Discovery of honeybee-inspired optic flow 'heuristics' by Srinivasan et al (1999) and Srinivasan and Zhang (2004) have spawned several approaches for local navigation (Santos-Victor and Sandini 1997; Franz and Mallot 2000; Serres et al 2008) and egomotion estimation (Franz and Krapp 2000; Franz et al 2004) (see also Chapter I,2 by MV Srinivasan et al.). Additionally, behavioral observations of optic flow expansion cues in fruit flies (Tammero and Dickinson 2002) have inspired reflexive obstacle avoidance strategies (Harrison 2005; Zufferey and Floreano 2006). While these efforts to transition behavioral heuristics to engineered systems have provided a path forward, the underlying principles are poorly understood; the results above have predominantly been presented without formal closed loop stability or performance analysis, and the approaches have largely ignored (in favor of more traditional architectures) the fundamental processing and feedback mechanisms that insects employ to extract information from optic flow and to regulate behavior.

2. Visuomotor feedback in insects

The insect retina, composed of thousands of individual sub-units, functions to image the incident patterns of luminance from the environment. As an insect moves, the intensity of the image formed at each lens becomes time dependent. The rate and direction of the local image shifts, taken over the entire field of view (FOV), form patterns of optic flow. The spatial structure of the patterns of optic flow that the insect experiences is governed primarily by the insect's relative motion and proximity to objects through motion parallax, a relationship that can be expressed mathematically in closed form (Koenderink and van Doorn 1997). Extraction of visual information contained in optic flow is performed by wide-field sensitive *tangential cells*, which communicate their output through descending neurons to the flight motor to execute changes in wing kinematics (Egelhaaf et al 2002; Borst and Haag 2002).

The tangential cells are large, motion-sensitive neurons that reside in the lobula plate portion of the visual ganglia (Fig. 1A). They are believed to integrate (pool) the outputs of large numbers of retinotopically-distributed elementary motion detectors (EMDs) (Franceschini et al 1989; Egelhaaf and Borst 1993; Krapp et al 1998; Egelhaaf et al 2002; Borst and Haag 2002). Prominent among the tangential cells are the identified neurons that comprise the 'horizontal system' (HS) and 'vertical system' (VS) found in a number of species of flies (Hausen 1982a,b; Hengstenberg et al 1982). As their names suggest, these neurons are sensitive primarily to horizontal and vertical patterns optic flow, respectively. They respond with graded membrane potentials whose polarity depends on the direction of motion. Their spatial sensitivity to local motion cues has in some cases been mapped out (Krapp et al 1998), as shown for the VS1 neuron in Fig. 1B, and the resemblance of some of these maps to the patterns of optic flow induced by particular modes of egomotion has led to the hypothesis that the corresponding neurons may act as matched filters for these patterns (Krapp and Hengstenberg 1996; Franz

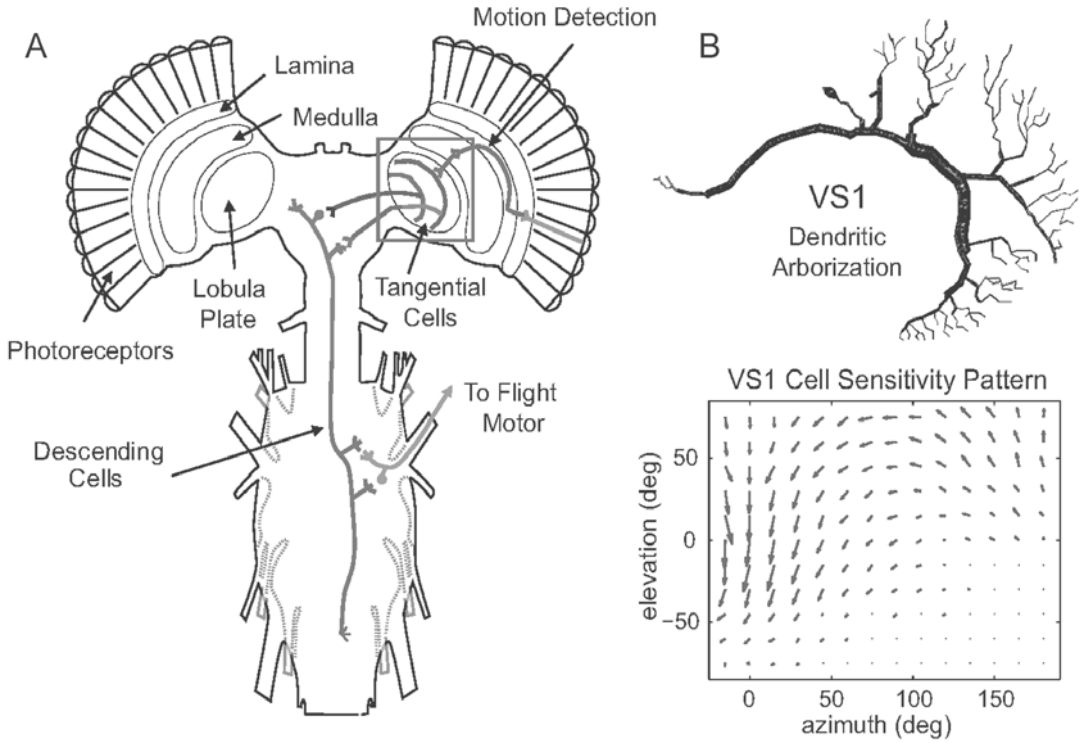


Fig. 1 **A** Visuomotor system structure. **B** Vertical system (VS1) tangential cell and associated wide-field directional sensitivity pattern. The data were extracted and replotted from Krapp et al (1998)

and Krapp 1998). However, recent work has shown that translational motion cues, which are the source of proximity information, are also present in the outputs of cells that were previously thought to be used only for compensation of rotary motion (Karmeier et al 2006). This suggests that cell patterns might be structured to extract a combination of relative speed and proximity cues, rather than direct estimates of the velocity state. Hence, while some progress has been made in understanding structure, arrangement, and synaptic connectivity (Egelhaaf et al 2002), the exact functional role that the tangential cells hold in the stabilization and navigation system of the fly remains a challenging and open question.

3. Wide-field processing of optic flow

In this section an inner product model for tangential cell analogues is presented and the Wide-Field Integration framework (Humbert and Hyslop 2010) is introduced to characterize the information that is encoded by the azimuthal pattern sensitivities of tangential cells that compose the horizontal system (HS). Small perturbation techniques are then applied to link weighting patterns to outputs which are functions of relative proximity and velocity with respect to the environment.

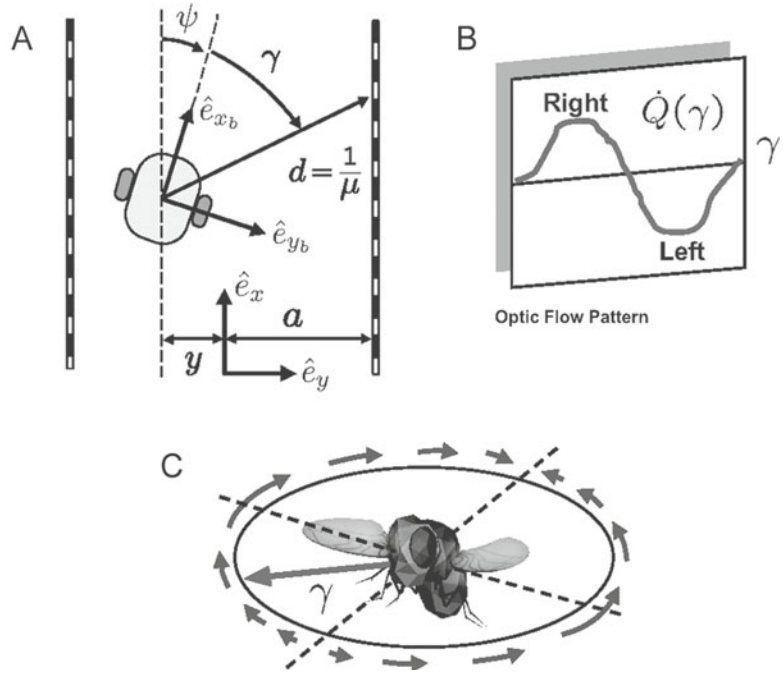


Fig. 2 **A** Environment approximation with planar vehicle, **B** nominal 1D flow as function of viewing angle, **C** nominal equatorial flow field around insect in an infinite tunnel

3.1 Optic flow model

The (true) optic flow is the vector field of relative velocities of material points in the environment projected into the tangent space of the imaging surface. It is a combination of the rotational and translational motion of the observer, along with the relative proximity to surrounding objects. For motion restricted to the equatorial plane, the optic flow on a circular imaging surface can be expressed as in Humbert and Hyslop (2010):

$$\dot{Q} = -\dot{\psi} + \mu\,(\dot{x}_b \sin\gamma - \dot{y}_b \cos\gamma), \qquad (1)$$

where ($\dot{x}_b$,$\dot{y}_b$) are the body-frame velocities, $\dot{\psi}$ is the yaw rate and μ is the nearness function which represents the distribution of objects in the surrounding environment and inherently encodes the vehicle's relative pose $\mathbf{q} = (x,y,\psi)$. Here ψ is defined as the heading angle, and x and y are the axial and lateral position in the tunnel, respectively (Fig. 2).

In order to completely specify the optic flow pattern (1) in closed form, simplifying assumptions are required on the shape of the nearness function $\mu(\gamma,\mathbf{q}) \in L_2[0,2\pi]$. The nearness is equal to $1/d(\gamma,\mathbf{q})$, where $d(\gamma,\mathbf{q}) \in (0,\infty)$ is the distance from the imaging surface to the nearest object in the environment at each viewing angle γ. For the analysis and the experiments described herein a planar corridor geometry is considered (Fig. 2), which provides an approximation of two wide-field obstacles. In this case the nearness function $\mu(\gamma,\mathbf{q})$ is independent of the axial position x and can be expressed in closed form as a function of the lateral position y, heading angle ψ, and the tunnel half-width a:

$$\mu(\gamma,\mathbf{q}) = \begin{cases} \frac{\sin(\gamma+\psi)}{a-y} & 0 \le \gamma+\psi < \pi \\ -\frac{\sin(\gamma+\psi)}{a+y} & \pi \le \gamma+\psi < 2\pi \end{cases} \qquad (2)$$

3.2 Tangential cell analogues

Tangential cells respond with graded membrane potentials or shifts in spiking frequency whose polarity depends on the direction

of motion. Essentially, the output is a *comparison* between a cell's preferred wide-field pattern of motion (e.g., Fig. 1B) and that of the visual stimulus. Mathematically, this comparison can be modeled as an *inner product* $\langle a,b \rangle$, analogous to the dot product between vectors, an abstraction of the *angle* between objects *a* and *b*. For planar optic flow about the yaw axis (Fig. 7), *tangential cell analogues* take the $L_2[0,2\pi]$ inner product form

$$\langle \dot{Q}, F \rangle = \frac{1}{\pi} \int_0^{2\pi} \dot{Q}(\gamma, \mathbf{x}) \cdot F(\gamma)\, d\gamma. \qquad (3)$$

Here γ is the body-referred viewing angle, $\dot{Q}(\gamma,\mathbf{x})$ is the measured optic flow pattern about the yaw axis, $F(\gamma)$ is any square integrable weighting function such that (3) exists, and $\mathbf{x} = (\dot{x}_b, \dot{y}_b, y, \theta, \dot{\psi})$ is defined as the state of the planar vehicle.

In the present study the analysis is restricted to motion in the horizontal plane, hence the sensitivity weightings for the H1, H2, Hx, HSN, HSE, HSS, dCH and vCH cells are considered for the left (L) and right (R) hemispheres. The restriction of optic flow measurements to the equatorial plane implies that only the azimuthalequatorial components of the sensitivity patterns are required for analysis. Representative data points were obtained from published data (Hausen 1982a,b; Krapp et al 1998) for each

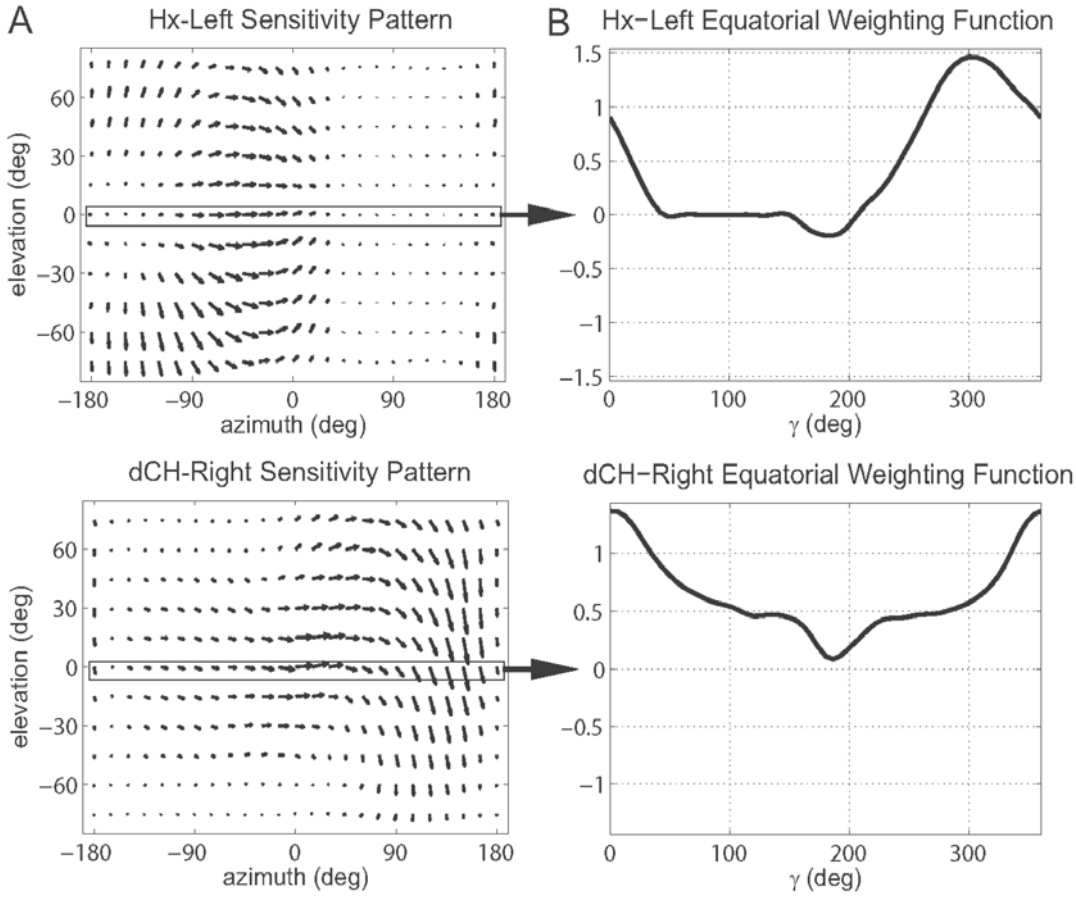


Fig. 3 Extraction of equatorial-azimuthal flow sensitivity for a left hemisphere horizontal system cell (Hx) and right hemisphere horizontal system (dCH) cell; **A** 2D optic flow sensitivity pattern, **B** azimuthal flow component for equatorial ring

cell (e.g., Fig. 3), smoothed with a Gaussian filter and converted to a continuous function of γ by making a 12th order Fourier series approximation. The resulting 16 weightings $\mathbf{F} = \{F_i, i = 1,\ldots,16\}$ are then normalized so that $\frac{1}{\pi}\int_0^{2\pi} F_i^2\, d\gamma = 1$. It follows that a perfect match of the flow field with the sensitivity weighting will result in a unity output.

To characterize the small signal content encoded by the azimuthal pattern sensitivity weightings, the set of outputs $\langle \dot{Q}, F_i \rangle$ corresponding to the collection of tangential cell weightings **F** from the left and right hemispheres are linearized about the nominal pattern of optic flow induced on the retina for centered motion within the environment parameterized by (2). This nominal pattern of optic flow corresponds to a equilibrium state of $\mathbf{x}_0 = \{v_0,0,0,0,0\}$. The resulting linearized observation equation $\mathbf{y} = C\mathbf{x}$ is given by

$$\begin{pmatrix} \mathrm{H1_R} \\ \mathrm{H1_L} \\ \mathrm{H2_R} \\ \mathrm{H2_L} \\ \mathrm{Hx_R} \\ \mathrm{Hx_L} \\ \mathrm{HSN_R} \\ \mathrm{HSN_L} \\ \mathrm{HSE_R} \\ \mathrm{HSE_L} \\ \mathrm{HSS_R} \\ \mathrm{HSS_L} \\ \mathrm{dCH_R} \\ \mathrm{dCH_L} \\ \mathrm{vCH_R} \\ \mathrm{vCH_L} \end{pmatrix} = \begin{pmatrix} -0.49 & 0.09 & -0.50 & -0.09 & 1.00 \\ -0.49 & -0.09 & 0.50 & 0.09 & -1.00 \\ -0.30 & 0.17 & -0.31 & -0.17 & 0.87 \\ -0.30 & -0.17 & 0.31 & 0.17 & -0.87 \\ -0.47 & 0.19 & -0.48 & -0.19 & 0.86 \\ -0.47 & -0.19 & 0.48 & 0.19 & -0.86 \\ 0.25 & -0.11 & 0.46 & 0.11 & -1.11 \\ 0.25 & 0.11 & -0.46 & -0.11 & 1.11 \\ 0.24 & -0.11 & 0.56 & 0.11 & -1.21 \\ 0.24 & 0.11 & -0.56 & -0.11 & 1.21 \\ 0.23 & -0.23 & 0.28 & 0.23 & -0.83 \\ 0.23 & 0.23 & -0.28 & -0.23 & 0.83 \\ 0.05 & -0.14 & 0.56 & 0.14 & -1.25 \\ 0.05 & 0.14 & -0.56 & -0.14 & 1.25 \\ 0.13 & -0.08 & 0.59 & 0.08 & -1.26 \\ 0.13 & 0.08 & -0.59 & -0.08 & 1.26 \end{pmatrix} \begin{pmatrix} \frac{1}{a}\dot{x}_b \\ \frac{1}{a}\dot{y}_b \\ \frac{v_0}{a^2}y \\ \frac{v_0}{a}\psi \\ \dot{\psi} \end{pmatrix}. \tag{4}$$

Inspection of (4) reveals that the output signals are functions of speed/depth quantities, due to the nature of optic flow. The generated outputs are highly coupled and no one cell weighting appears to provide direct measurement of any particular speed/depth quantity as they have been parameterized. However, it is clear that a forward speed estimate $\dot{x}_b$ can be obtained by summing the outputs of two same-type cells from opposite hemispheres due to their symmetric sensitivity.

4. Closed loop feedback analysis and synthesis

In the following section the feedback interconnection shown in Fig. 4 is analyzed to develop a methodology to link tangential cell weighting patterns to reflexive navigation behavior. Mathematical tools from control and estimation theory are introduced that separate the task of tangential cell feedback gain selection into two steps; (a) static state estimation, and (b) state feedback gain synthesis.

4.1 Extraction of relative state information

Optic flow cannot be measured directly, it must be inferred from the spatiotemporal patterns of luminance incident on an imaging surface or retina. Therefore, the optic flow estimation process introduces error in the measurements, which is compounded by sensor noise and contrast/texture variations occurring in the environment. Given $m \geq n$ linearly independent tangential cell weighting functions $\mathbf{F} = \{F_i, i = 1,\ldots,m\}$ and accounting for measurement error, the observation equation (4) can be expressed as

$$\tilde{\mathbf{y}} = C\mathbf{x} + \mathbf{w}, \tag{5}$$

where $\tilde{\mathbf{y}} \in \mathbb{R}^m$ are the measured outputs and w is the measurement error.

The static estimation objective is to select an inversion matrix $C^{\dagger} \in \mathbb{R}^{n\times m}$ that provides estimates $\hat{\mathbf{x}}$ that minimize the error between the measurements (5) and the model (4) for a given set of weighting functions **F**. The problem is posed as a standard static linear estimation problem, where one seeks the solution of an overdetermined ($m \geq n$), inconsistent set of linear equations given by (5). The optimal choice that minimizes the

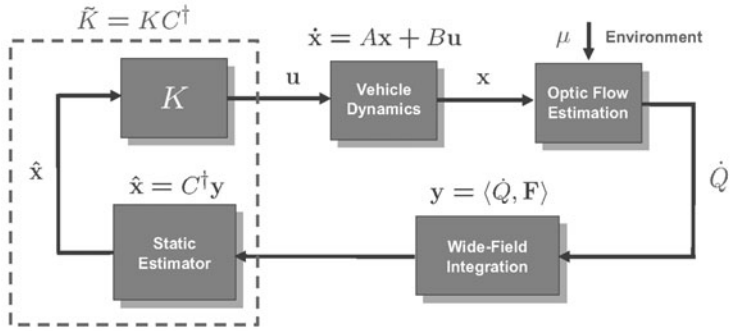


Fig. 4 Visuomotor system analogue. On-board optic flow measurements $\dot{Q}$ are decomposed using tangential cell weightings **F** into outputs **y**. Feedback gains $\tilde{K} = KC^{\dagger}$ are applied to tangential cell outputs to generate closed loop feedback commands **u**

sum of squares of the residual errors $J = \frac{1}{2}(\tilde{\mathbf{y}} - C\hat{\mathbf{x}})^T(\tilde{\mathbf{y}} - C\hat{\mathbf{x}})$ is given by the *least squares (LS) estimator* $\hat{\mathbf{x}} = C^{\dagger}\tilde{\mathbf{y}}$, where

$$C^{\dagger} = \left(C^T C\right)^{-1} C^T. \tag{6}$$

If w is assumed to be a random variable with known mean $E\{\mathbf{w}\} = 0$ and covariance $E\{\mathbf{w}\mathbf{w}^T\} = \mathbf{R}_w$, the Gauss-Markov theorem yields the *minimum variance (MV) estimator*,

$$C^{\dagger} = (C^T R_{\mathbf{w}}^{-1} C)^{-1} C^T R_{\mathbf{w}}^{-1}. \tag{7}$$

To obtain an expression for R_w in the context of tangential cell analogues, it is first assumed that optic flow measurements taken at discrete locations on the circle (with equal spacing $\Delta\gamma$) are affected by zero mean additive noise $\eta(\gamma)$ with variance σ_η^2, and no correlation between measurement nodes or with signal magnitude. Under these assumptions the measurement noise at the level of the integrated output is $w_i = \langle \eta, F_i \rangle$. Using linearity of (3) and the properties of the covariance matrix one obtains

$$R_{\mathbf{w},ij} = \frac{\Delta\gamma\,\sigma_\eta^2}{\pi}\,\langle F_i, F_j \rangle. \tag{8}$$

Note that the wide-field integrated measurement noise approaches zero as the number of measurement nodes on the imager approaches infinity, providing significant improvement in signal to noise ratio -an attractive property of the wide-field integration processing approach. Additional conclusions relating noise throughput performance of the estimators can be quantified using Fisher information, as shown in Hyslop et al (2010).

Now consider relative state estimates $\hat{\mathbf{x}} = C^{\dagger}\mathbf{y}$, where $\mathbf{y} = \langle \dot{\mathbf{Q}}, \mathbf{F} \rangle$. If the inversion is pushed through the inner product, one obtains

$$\hat{x}_i = \left\langle \dot{\mathbf{Q}}\,,\, \sum_{j=1}^{m} C_{ij}^{\dagger} F_j \right\rangle, \tag{9}$$

where the second argument in the inner product can be interpreted as the optimal extraction pattern for the ith state, $F_{\hat{x}_i} = \sum_{j=1}^{m} C_{ij}^{\dagger} F_j$. This can be interpreted as the relative weights one would apply to the instantaneous optic flow measurements to generate an estimate of a particular state. Hence, the set of optimal state extraction patterns (Fig. 5), are given by

$$\mathbf{F}_{\hat{\mathbf{x}}} = C^{\dagger}\mathbf{F}. \tag{10}$$

When a vehicle state becomes non-zero it imposes a perturbation pattern upon the nominal optic flow field. These patterns are not necessarily orthogonal to other state-induced patterns, thus the optimal inversion generates weighting functions $\mathbf{F}_{\hat{\mathbf{x}}}$ that can be interpreted as removing the coupling due to non-orthogonal state perturbations.

There is one important feature of the observation equation (4) that warrants further discussion here. Note that even though the vehicle states were defined relative to the

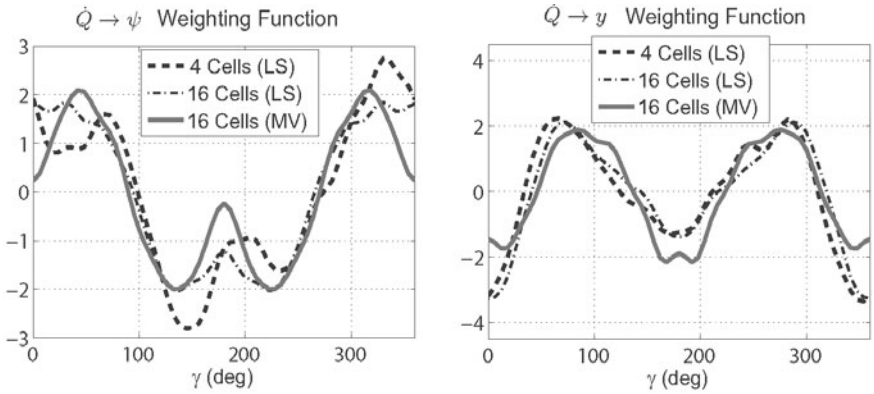


Fig. 5 Optimal state extraction pattern $\mathbf{F}_{\hat{\mathbf{x}}} = C^{\dagger}\mathbf{F}$ comparison for three different tangential cell weighting function set selections; LS = least squares estimator, MV = minimum variance estimator

plane of symmetry of the insect, the directional motion preferences of the cells do not appear to be tuned to the natural forward/lateral motions as one would see in a traditional engineered system, i.e., the C matrix is not diagonal. Furthermore, an individual cell's output depends on not just several, but the entire set of proximity and velocity states, which significantly obfuscates the feedback gain selection process. As will be demonstrated in Section 5, the optimal state extraction procedure described in this section provides a way to separate these dependencies, allowing for the application of traditional tools from control theory for principled design of feedback gains.

4.2 Static output feedback

Assuming available estimates $\hat{\mathbf{x}}$ of the vehicle state, any number of linear feedback design techniques (Stevens and Lewis 2003) can be used to synthesize gains K that provide a stabilizing feedback

$$\mathbf{u} = K\hat{\mathbf{x}}, \tag{11}$$

given the vehicle dynamics model. It is clear from Fig. 4 that the resulting output feedback gains that are applied at the level of tangential cell outputs to generate the appropriate input are given by $\tilde{K} = KC^{\dagger}$. Therefore, what has been shown is that the gain selection task for a given set of tangential cell weighting patterns can be cast in a rigorous framework in which traditional tools from control and information theory can be leveraged to achieve the desired closed loop visual-based behaviors. Section 5.1 discusses the application and validation of this methodology on a wheeled robot.

5. Experimental validation

In the following section the methodology for achieving reflexive navigation behavior based on tangential cell outputs is demonstrated on a wheeled vehicle. Azimuthal optic flow is computed on board using an implementation of the Lucas-Kanade pyramid iterative algorithm (Lucas and Kanade 1981) from imagery obtained from a camera pointed upward at a parabolic mirror which provides a 360° field of view of the surround-

ing environment. Tangential cell weighting functions are applied to optic flow estimates obtained at $k = 20$ discrete locations around the azimuth of the vehicle, and the angular velocity command is generated at 55 Hz and applied to the wheel speed controllers. For a detailed system level description of the hardware and software implementation the reader is referred to Humbert and Hyslop (2010).

5.1 Feedback synthesis

A kinematic model for the motion of the wheeled robot with a constant forward velocity $\dot{x}_b = v_0$ is given by

$$\begin{aligned} \dot{y} &= v_0 \sin\theta \\ \dot{\psi} &= u, \end{aligned} \tag{12}$$

where y is the lateral position and ψ is the heading. The angular velocity input $u \in \mathbb{R}$ is intended to generate commands to reflexively maneuver the vehicle between objects in the surrounding environment. The output feedback is implemented as

$$u = \tilde{K}\tilde{\mathbf{y}}, \tag{13}$$

where the control u is generated by applying gains $\tilde{K} = KC^{\dagger}$ to outputs

$$\tilde{y}_i(t) = \frac{\Delta\gamma}{\pi} \sum_{j=1}^{k} \dot{Q}(\gamma_j, t)\, F_i(\gamma_j), \tag{14}$$

which are computed from the instantaneous measurements of optic flow $\dot{Q}(\gamma_j, t)$ and the set $\mathbf{F} = \{F_i(\gamma_j),\ i = 1,\ldots,m\}$ of tangential cell weightings.

To select the associated state feedback gains $K = [K_\psi\ K_y]$ we assume estimates $\hat{\mathbf{x}} = [\hat{\psi}\ \hat{y}]^T$ that are used to generate the closed loop command $u = K_\psi\hat{\psi} + K_y\hat{y}$. The resulting linearized closed loop dynamics are

$$\begin{pmatrix} \dot{y} \\ \dot{\theta} \end{pmatrix} = \begin{pmatrix} 0 & v_0 \\ K_y & K_\psi \end{pmatrix} \begin{pmatrix} y \\ \theta \end{pmatrix}, \tag{15}$$

with associated characteristic equation $s^2 - K_\psi s - v_0 K_y = 0$ and eigenvalues

$$s_{1,2} = \frac{1}{2}\left(K_\psi \pm \sqrt{K_{\psi^2} + 4v_0 K_y}\right). \tag{16}$$

Clearly it is required that $K_\psi < 0$ and $K_y < 0$ for stability; additional performance criteria such as overshoot and rise time can be imposed by selecting the ratio of K_ψ relative to K_y.

To determine the inversion $C^{\dagger}$, several sets of weightings were considered for comparison: (a) a subset of 4 tangential cell weighting patterns (HX_R, HX_L, HSS_L and VCH_R) using the least-squares inversion (6); (b) all 16 tangential cells using the least squares inversion (6); and (c) all 16 tangential cells using the minimum variance inversion (7). The resulting state extraction patterns $\mathbf{F}_{\hat{\mathbf{x}}} = C^{\dagger}\mathbf{F}$ are plotted in Fig. 5. The 4 cell case represents the minimum number of TCs required to guarantee C is full rank, and hence existence and uniqueness of a solution to (5). Note that the sideslip velocity $\dot{y}_b = 0$ for a wheeled vehicle. If this were not the case then outofplane optic flow measurements and a patterns with out-of-plane sensitivity would be required to maintain full rank of the measurement model (Hyslop and Humbert 2010).

5.2 Results

The closed loop implementation was tested in two environments; a corridor with a 90° bend (Fig. 6A), and a cluttered field of large obstacles (Fig. 6B). The ratio of state feedback gains $K_y : K_\psi$ for the trials shown was selected to be 2.25 : 1 to balance closed loop overshoot and rise time. In the latter environment two different initial conditions are used to evaluate performance. The vehicle's trajectory was recorded using a VICON tracking system, which employs 8 cameras to triangulate the position of reflective markers attached to the vehicle. Measurements

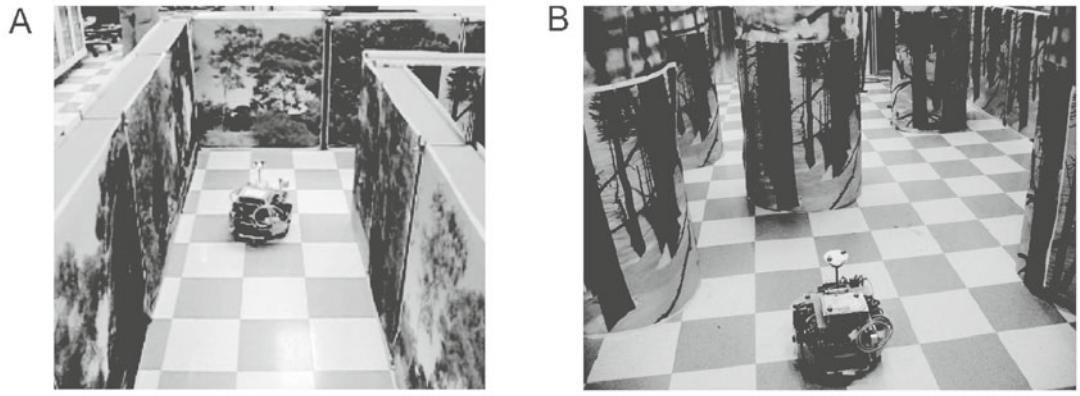


Fig. 6 **A** Corridor environment with 90° bend. **B** Cluttered obstacle field environment

were recorded at 350 Hz and is accurate to less than 1 mm, and trials were repeated 10 times.

Figure 7A–C shows that the vehicle successfully navigates the tunnel environment using all three inversions. The only clear trend is that 4-cell LS feedback avoids overshoot during the initial condition recovery and executes a sharper turn at the 90° bend. The reason is that the 4-cell feedback is fundamentally biased toward detecting off-nominal flow on the front-left and rear-right of the vehicle (see the ψ sensitivity function asymmetry in Fig. 5B). During the initial recovery, the presence of the 90° bend downrange causes reduced optic flow on the vehicle's front-right. With the reduced sensitivity in this region, the 4-cell LS feedback does not respond to the optic flow perturbation and reacts as if it were in a straight corridor. Implementation of the 16-cell feedback (symmetric state sensitivity functions) demonstrates that the optic flow perturbation leads to a reduced ψ-estimate which causes overshoot in the centering behavior.

Results for the cluttered object field environment are shown in Fig. 7D–F. The bias in the 4-cell feedback is evident in the initial condition starting at y = 0.3 m (Fig. 7D), but with adverse effects. Due to this bias, the vehicle fails to avoid the final right-side cylinder in time, resulting in impact in 9 of the 10 trials. This underscores the importance of including symmetric pairs of weighting functions in the decomposition of optic flow

Table 1 Tangential cell feedback gains $\tilde{K}$ for rotation rate control, using Fig. 4 control loop architecture

Cell Pattern	4-Cell LS	16-Cell LS	16-Cell MV
$H1_R$		42	371
$H1_L$		–42	–371
$H2_R$		–66	–40
$H2_L$		66	40
Hx_R	–287	226	212
Hx_L	229	–226	–212
HSN_R		78	–22
HSN_L		–78	22
HSE_R		9	–167
HSE_L		–9	167
HSS_R		35	–84
HSS_L	–232	–35	84
DCH_R		11	729
DCH_L		–11	–729
VCH_R	199	32	–73
VCH_L		–32	73

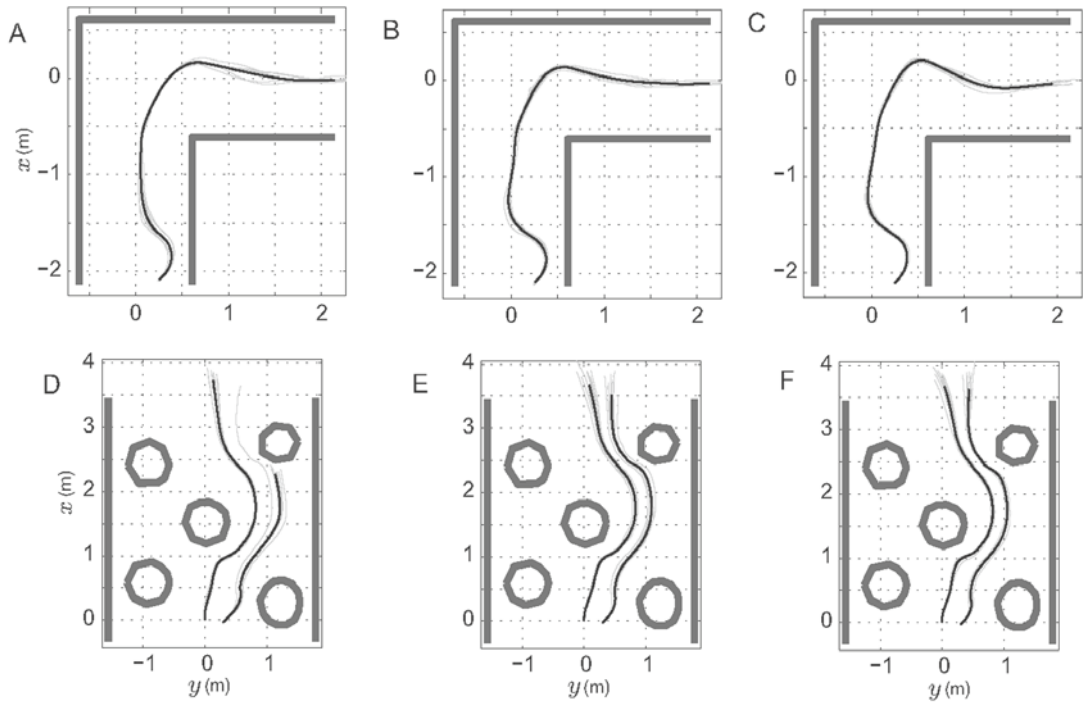


Fig. 7 Vehicle trajectories (10 trials) and mean trajectory for tunnel environment with 90° bend and cluttered obstacle field (forward speed v_0 = 0.4 m/s); tangential cell gains determined from **A**, **D** 4-cell LS, **B**, **F** 16-cell LS, and **C**, **F** 16-cell MV

patterns to realize a symmetric mapping between optic flow measurements and actuator response.

Discussion

The mathematical analogue presented in this chapter demonstrates how stabilizing feedback gains can be synthesized for any set of *m* linearly independent tangential cell weighting patterns, providing that $m > n$ where *n* is the number of states required for closed loop vehicle control. Furthermore, the patterns required to achieve specific visual-based navigation are not unique; similar behaviors can be achieved with quite different sets of weighting functions (Table 1). Hence, the only requirement on selection of weighting function sets is that they are sufficiently different (independent) from one another (a large collective span). This is closely related to recent results from the field of signal compression and reconstruction, where the most efficient method of signal encoding is to use random basis functions if one has no prior model for the signal.

As evidenced by Fig. 7 D–F, the achievable navigation behaviors using feedback of tangential cell outputs are not limited to navigation of corridors. The internal state measurements (orientation with respect to obstacles and lateral offset from a centered path) simply represent asymmetries in the nominal optic flow pattern or perturbations from the equilibrium (Taylor and Krapp 2008; Humbert and Hyslop 2010). Regulation of these quantities to zero allows the vehicle to maintain a symmetric optic flow pattern, the consequence of which is a safe, centered trajectory between large obstacles. This is precisely the navigational heuristic observed in honeybees in Srinivasan et al (1996). If different closed loop heuristic behaviors are desired, such as wall following, they can be rapidly realized by adjusting the nominal optic flow pattern. It is important to note, however, that the methodology presented here is restricted to the avoidance of medium to large obstacles. For detection of small objects, insects make use of dif-

ferent small-field processing mechanisms (Egelhaaf 1985). In addition, the analysis presented in this chapter assumes planar movement and small motions about the linearized equilibrium. The impact of nonlinear dependencies associated with perturbations far from equilibrium have been addressed in separate work (Humbert and Hyslop 2010), where the full nonlinear stability of the underlying feedback methodology was proven using Lyapunov techniques. In addition, the conclusions presented in this work hold and for six degree-of-freedom (non-planar) motion as shown in separate work (Hyslop and Humbert 2010).

The findings in Section 4 help explain results from similar studies involving use of tangential cells for estimation and control. State estimation was investigated by Karmeier et al (2006), who suggested using a summation of same-type horizontal cell outputs for forward speed estimation, and subtraction of same-type outputs for yaw rate or sideslip. These claims are supported by inspection of (4). However, in Karmeier et al (2006) the results were limited to unweighted two-cell signal combinations, which essentially constitutes inversion of a non-full rank measurement set. No method is proposed for decoupling yaw rate from sideslip, or from the proximity and orientation states that are also embedded in the subtraction-method signal. The visuomotor analogue presented here provides a rigorous methodology for using tangential cell weighting properties in closed loop feedback.

Closed-loop control was attempted in Lindemann et al (2008), where obstacle avoidance was attempted in simulation using a comparison of the left and right HSE cell outputs. While this was useful for obstacle detection during periods of non-rotation, it was unsuccessful when applied as continuous feedback because the subtracted HSE signal contains coupled measurements of multiple states. The actuator signal should indeed include these quantities (yaw rate, sideslip, proximity and orientation), but they need to be fused using appropriate stabilizing gains rather than in the inherent ratios of the HSE pattern. Although in Lindemann et al (2008) a hi-fidelity neuronal-based model for the wide-field integration was used, the primary information throughput remains the same.

Summary

This chapter provides an analogue to the insect visuomotor processing, specifically tangential cells from the HS system, which can be directly applied to robotic platforms for robust obstacle avoidance and stability augmentation. The analogue is not an attempt to precisely model the exact feedback interconnections within insects, but it does seek to distil the fundamental operational principles and characterize them within a framework based on the appropriate mathematics. Specifically, it is shown that the resulting feedback synthesis task can be cast as a combined static state estimation and linear feedback control problem. The framework described above additionally allows for the analysis and determination of direct mapping between optic flow measurements and actuator commands, which greatly simplifies implementation on robotic platforms.

Acknowledgements

The authors would like to thank Alex Brown for extracting the tangential cell sensitivity data from references Hausen (1982a,b); Krapp et al (1998).

References

Barrows GL, Chahl J, Srinivasan M (2003) Biologically inspired visual sensing and flight control. Aeronaut J 107:159–168 Borst A, Haag J (2002) Neural networks in the cockpit of the fly. J Comp Physiol A 188: 419–437

Egelhaaf M (1985) On the neuronal basis of figure-ground discrimination by relative motion in the visual system of the fly. ii. figure detection cells, a new class of visual interneurones. Biol Cybern 52: 195–209

Egelhaaf M, Borst A (1993) Motion computation and visual orientation in flies. J Comp Biochem Physiol A 104A: 659–673

Egelhaaf M, Kern R, Krapp H, Kretzberg J, Kurtz R, Warzecha A (2002) Neural encoding of behaviourally relevant visual-motion information in the fly. Trends Neurosci 25: 96–102

Franceschini N, Riehle A, Nestour AL (1989) Directionally selective motion detection by insect

neurons. In: Stavenga D, Hardie R (eds) Facets of vision, Springer Verlag, Berlin, pp 360–390

Franz MO, Krapp HG (1998) Wide-field, motion-sensitive neurons and matched filters for optic flow fields. Biol Cybern 83: 185–197

Franz MO, Krapp HG (2000) Wide-field, motion-sensitive neurons and matched filters for optic flow fields. Biol Cybern 83: 185–197

Franz MO, Mallot HA (2000) Biomimetic robot navigation. J Robot Auton Syst 30: 133–153

Franz MO, Chahl JS, Krapp HG (2004) Insect-inspired estimation of egomotion. Neural Comput 16: 2245–2260

Gibson JJ (1950) The perception of the visual world. Houghton Mifflin, Boston

Harrison R (2005) A biologically inspired analog ic for visual collision detection. IEEE Trans Circuits Syst I 52(11):2308–2318

Hausen K (1982 a) Motion sensitive interneurons in the optomotor system of the fly, part i. the horizontal cells: Structure and signals. Biol Cybern 45: 143–156

Hausen K (1982 b) Motion sensitive interneurons in the optomotor system of the fly, part ii. the horizontal cells: Receptive field organization and response characteristics. Biol Cybern 46: 67–79

Hengstenberg R, Hausen K, Hengstenberg B (1982) The number and structure of giant vertical cells (vs) in the lobula plate of the blowfly *Calliphora Erythrocephala*. J of Comp Physiol 149: 163–177

Humbert JS, Hyslop AM (2010) Bioinspired visuomotor convergence. IEEE Trans Robot 26(1):121–130

Hyslop A, Humbert J (2010) Autonomous navigation in 3-d urban environments using wide-field integration of optic flow. J Guid Cont Dyn 33(1):147–159

Hyslop AM, Krapp HG, Humbert JS (2010) Control theoretic interpretation of directional motion preferences in optic flow processing interneurons. Biol Cybern 103(5):339–352

Karmeier K, van Hateren J, Kern R, Egelhaaf M (2006) Encoding of naturalistic optic flow by a population of blowfly motion-sensitive neurons. J Neurophysiol 96: 1602–1614

Koenderink JJ, van Doorn AJ (1997) Facts on optic flow. Biol Cybern 56: 247–254

Krapp HG, Hengstenberg R (1996) Estimation of self-motion by optic flow processing in single visual interneurons. Letters to Nature 384: 463–466

Krapp HG, Hengstenberg B, Hengstenberg R (1998) Dendritic structure and receptive-field organization of optic flow processing interneurons in the fly. J Neurophysiol 79: 1902–1917

Lindemann JP, Weiss H, Mller R, Egelhaaf M (2008) Saccadic flight strategy facilitates collision avoidance: closed-loop performance of a cyberfly. Biol Cybern 98: 213–227

Lucas B, Kanade T (1981) An iterative image registration technique with an application to stereo vision. In: Proceedings of the 7th International Joint Conference of Artificial Intelligence

Santos-Victor J, Sandini G (1997) Embedded visual behaviors for navigation. J Robot Auton Syst 19: 299–313

Serres J, Dray D, Ruffier F, Franceschini N (2008) A vision-based autopilot for a miniature air vehicle: joint speed control and lateral obstacle avoidance. Auton Robot 25: 103–122

Srinivasan M, Chahl JS, Weber K, Venkatesh S, Nagle MG, Zhang SW (1999) Robot navigation inspired by principles of insect vision. J Robot Auton Syst 26: 203–216

Srinivasan MV, Zhang SW (2004) Visual motor computations in insects. Ann Rev of Neurosci 27: 679–696

Srinivasan MV, Zhang SW, Lehrer M, Collet TS (1996) Honeybee navigation *en route* to the goal: Visual flight control and odometry. J Exp Biol 199: 237–244

Stevens B, Lewis FL (2003) Aircraft control and simulation. John Wiley & Sons, Inc., Hoboken, NJ

Tammero LF, Dickinson MH (2002) The influence of visual landscape on the free flight behavior of the fruit fly *Drosophila melanogaster*. J Exp Biol 205: 327–343

Taylor G, Krapp H (2008) Sensory systems and flight stability: What do insects measure and why? Adv Insect Physiol 34: 231–316

Wood RJ, Avadhanula S, Sahai R, Steltz E, Fearing RS (2008) First takeoff of a biologially-inspired at-scale robotic insect. IEEE Trans Robot 24(2): 341–347

Zufferey JC, Floreano D (2006) Fly-inspired visual steering of an ultralight indoor aircraft. IEEE Trans Robot 22: 137–146

Biologically inspired enhancement of dim light video

5

Henrik Malm, Magnus Oskarsson, Eric Warrant

Contents

Abstract

In this chapter a technology for the enhancement of video data obtained at low light levels is presented. The method was inspired by the way in which nocturnal animals adaptively sum intensities, spatially and temporally, to improve vision at night. Due to the low photon count under these conditions the visual input is dark and unreliable, which leads to noisy low contrast images. The noise becomes very apparent when we try to enhance the contrast and, by this, amplify the intensities in the darkest regions of the images. By constructing spatio-temporal smoothing kernels that automatically adapt to the three dimensional intensity structure at every point, the noise can be considerably reduced, with fine spatial detail being preserved and enhanced without added motion blur. For color image data, the chromaticity is restored and demosaicing of raw RGB input data can be performed simultaneously with the noise reduction. The method is a very generally applicable one, contains only few user-defined parameters and has been developed for efficient parallel computation using a graphics processing unit (GPU). The technique has been applied to image sequences with various degrees of darkness and noise levels. Results from some of these tests, and comparisons to related work, are presented here.

Henrik Malm
Lund University, Department of Cell and Organism Biology, Zoology Building
Helgonavägen 3, 22362 Lund, Sweden
e-mail: henrik.malm@cob.lu.se

1. Introduction

Cameras have the same problems at low light levels as we humans do; a low photon count makes the visual input dark and unreliable and this leads to noisy, low-contrast,

images. We have all experienced these difficulties when trying to capture a special moment with a digital camera or mobile phone when the lighting is poor. These problems are even more severe in video than in still images because of temporal fluctuations, which can be seen when trying to amplify the intensities.

Furthermore, the dynamic range of a standard digital camera is very limited, relying on a single exposure time and a uniform sensitivity across the imaging sensor. If light intensities in the scene being captured span a wide range, this inevitably leads to either under- or overexposed areas in the images. If we simply try to amplify dark under-exposed areas, uniformly or by a non-linear intensity transformation, the low signal-to-noise ratio in those areas will become very apparent. Apart from the uncorrelated Poisson-distributed variations arising from the low photon count, the so-called photon shot noise, there are several other sources of noise in a camera, including read-out noise, pattern noise, quantization errors and temperature induced noise (Reibel et al. 2003; Alter et al. 2006).

In order to understand how to deal with the above-mentioned problems and how to increase the range of usability of ordinary cameras under conditions of low light, we have turned to nocturnal animals and their remarkable visual systems for primary inspiration.

1.1 Biological inspiration

Surviving in habitats with high competition for resources and the threat of predation, many animals have evolved a nocturnal lifestyle. To be able to navigate and forage during the hours of darkness, these animals have developed excellent vision at light levels where humans are practically blind, even though many species have eyes and brains that are orders of magnitude smaller than those of humans (Warrant 2004, 2008). It is therefore interesting to study the visual systems of these animals to gain insight into the mechanisms that permit these impressive visual capabilities.

A characteristic of the visual systems of nocturnal animals, both vertebrates and invertebrates, is their highly developed ability to sum the visual signal locally in space and time to increase signal strength and in this way improve the reliability of intensity estimations (Warrant 1999, 2008). In the temporal domain, when light gets dim, the visual systems of nocturnal animals can integrate signals over longer periods of time (van Hateren 1993, Warrant 1999). In the eye, this can be achieved by having slower photoreceptors, as is typical of nocturnal insects (Laughlin and Weckström 1993, Frederiksen et al. 2008). Even slower vision could be obtained by neurally integrating (summing) signals at a higher level of integration in the visual system. In the spatial domain, the transition to dim light could activate specialised laterally spreading neurons which couple the channels together into groups (van Hateren 1993, Warrant 1999). Each summed group – themselves now defining the channels – could collect considerably more photons over a much wider visual angle than a single channel.

These spatio-temporal summation mechanisms are adaptive and become more pronounced at lower light levels, but their benefits come at the cost of lower spatial and temporal resolution. The relative amount of spatial and temporal summation depends on the specific needs and lifestyle of the animal in question. Nocturnal animals that move fast and need to react to high speed events (e. g. flying moths) require good temporal resolution, which therefore favours spatial summation over temporal summation. In contrast, in nocturnal animals that move slowly, temporal summation is favoured in order to maximize spatial resolution. War-

rant (1999) constructed a model for visual signal processing in animals that can be used to find the optimal extents of spatial and temporal summation that maximize visual performance at a given light intensity and image velocity. This biological mechanism of adaptive spatio-temporal smoothing has been the initial inspiration for the image processing technique presented here.

We have, during our work, noticed parallels between our complete system and the combined mechanisms by which nocturnal animals process visual information at night. Our image processing approach for dim light vision also includes an initial step for contrast enhancement, in order to enhance the visibility of structures in the darker areas of the images, and a mechanism for sharpening coarser edges in the scene. The contrast enhancement step is, for example, akin to the increased amplification typically found in nocturnal insect photoreceptors (Laughlin et al. 1993, Frederiksen et al. 2008). The photoreceptors of the nocturnal bee *Megalopta genalis* have a gain up to 5 times higher than that found in the photoreceptors of closely related day-active bees (Frederiksen et al. 2008). Unfortunately, the increased amplification not only affects the signal but also the noise, and thus on its own, the higher gain does not alter the visual signal-to-noise ratio. However, by subsequently applying spatio-temporal summation, the noise can be efficiently averaged out, thereby greatly improving the signal-to-noise ratio. For edge sharpening, animal visual systems typically employ neural mechanisms – often based on lateral inhibitory networks (Hartline and Ratliff 1957; van Hateren 1992, 1993). Using these mechanisms edges in the visual scene can be enhanced neurally so that they are more reliably detected by the visual system. This sharpening of the coarser details, which remain following the signal summation processes, is likely to help maximize visual reliability. Lateral inhibition in the animal visual system has a direct parallel to the sharpening feature of our image processing technology.

1.2 Image processing overview

The first step in our image processing approach is contrast enhancement. Here we have chosen to apply so-called *contrast-limited histogram equalization* (Pizer et al. 1987), which will be presented in detail in Section 3. After contrast enhancement, the noise in the image sequence will become very apparent, especially in the formerly dark areas of the images. Inspired by insect vision, we apply spatio-temporal summation. However, instead of applying a combination of pure spatial and pure temporal summation, as done in the biological model of Warrant (1999), we treat the three dimensions in the spatio-temporal space equally. To this end, a three-dimensional spatio-temporal function is constructed at every pixel, which is then used to calculate a weighted average around this pixel. By doing so, we efficiently smooth the signal and reduce the noise, while at the same time preserving important object contours and preventing motion blur. In order to counteract any blurring of edges that still might appear, the spatio-temporal weighting function can be modified to also include a sharpening feature. The noise reduction step, including the sharpening extension, is covered in detail in Section 4.

After this presentation of the core theory, we will discuss how the algorithm can be used to process colour image data (see Section 5), and how the algorithm can be implemented using the parallel computational power of a graphics processing unit (GPU) on a standard graphics card (see Section 6). We will end our presentation by reviewing some experimental results (see Section 7). First, however, we review related work in the literature.

2. Related work

There is a multitude of noise reduction techniques that apply spatio-temporal weighted averaging for noise reduction purposes. Many authors have additionally realized the benefit of first estimating the motion in the image sequence and then using this information in the spatio-temporal filtering. If the motion is known, the spatio-temporal filters could be orientated in the direction of the motion in the spatio-temporal space. This will reduce the amount of added motion blur and the amount of spatial averaging needed. These noise reduction techniques are usually referred to as motion-compensated spatio-temporal filtering. Kalivas and Sawchuk (1990) calculated means along the motion trajectories while Özkan et al. (1993) and Miyata and Taguchi (2002) apply weighted averages, dependent on the intensity structure in a small neighbourhood. Sezan et al. (1991) used so-called linear minimum mean-square-error filtering along the trajectories.

Most of the motion-compensated methods use some kind of block-matching technique for the motion estimation that usually, for efficiency, employs rather few images, commonly 2 or 3 in total for each filtered image. Several variants using optical flow estimators do however exist (Sezan et al. 1991, Özkan et al. 1993; Dugad and Ahuja 1999). A drawback of motion-compensating methods is that the filtering relies on a good motion estimation to give a good output without excessive blurring. The motion estimation is especially complicated for sequences severely degraded by noise. Various approaches have been applied to deal with this problem, often by simply reducing the influence of the temporal filter in areas where the reliability of the motion estimation is low. This, however, frequently entails disturbing noise at object contours, for example.

Another class of noise-reducing video processing methods uses a cascade of directional filters and in this way analyzes the spatio-temporal intensity structure in the neighbourhood of each point. The filtering and smoothing is done primarily in the direction that corresponds to the filter with the highest output (Martinez and Lim 1985; Arce 1991; Ko and Forest 1993). These methods work well for directions that coincide with the fixed filter directions but the output is considerably degraded if the actual direction of motion lies between the filter directions. For a review of spatio-temporal noise reduction methods see Brailean et al. (1995).

An interesting family of smoothing techniques for noise reduction are the ones that solve an edge-preserving anisotropic diffusion equation on the images. In this way, the intensities are treated as if they were concentrations of a certain substance and the smoothing comes from the physical equalization of these concentrations if the substance is allowed to diffuse. This approach was pioneered by Perona and Malik (1990) and has had many successors, including the work by Weickert (1998). It has also been extended to 3D and spatio-temporal noise reduction in video by Uttenweiler et al. (2003) and Lee and Kang (1998). Uttenweiler et al. (2003) and Weickert (1999) apply the so-called *structure tensor* or *second-moment matrix* in a manner similar to our approach in order to analyze the local spatio-temporal intensity structure and steer the smoothing accordingly. The drawbacks of techniques based on diffusion equations include the fact that finding the solution needs a time-consuming iterative procedure. Moreover, it is very difficult to find a suitable stopping time for the diffusion, at least in an automatic manner. These drawbacks often make these approaches unsuitable for video processing.

A better approach is to apply single-step structure-sensitive adaptive smoothing kernels. The bilateral filters introduced by To-

masi and Manduchi (1998) for 2 D images fall within this class. Here, the smoothing is performed by calculating a weighted average at every point using a Gaussian kernel, where the coefficients in the kernel are attenuated based on how different the intensities are in the corresponding pixels compared to the centre pixel. In this way, the edges are maintained since no averages are calculated across sharp edges. However, relating all coefficients to the central pixel makes the local smoothing very dependent on the correctness of the intensity in this pixel, whose value cannot be easily estimated in images heavily disturbed by noise. This drawback can be somewhat reduced by a pre-smoothing before pixel comparisons, but inevitably this also leads to a higher degree of over-all smoothing. A comparison with the median intensity in a local neighbourhood around the central pixel has also been suggested.

An approach closely connected to both bilateral filtering and to anisotropic diffusion, techniques based on the structure tensor, is the structure-adaptive anisotropic filtering (Yang et al. 1996). This is the work that most closely resembles our approach and we expand on this in Section 4. For a study of the connection between anisotropic diffusion, adaptive smoothing and bilateral filtering, see Barash and Comaniciu (2004).

Another group of algorithms connected to our work is in the field of *high dynamic range* (HDR) imaging (Kang et al. 2003; Reinhard et al. 2005). The aim of HDR imaging is to alleviate the restriction caused by the low dynamic range of ordinary CCD cameras, i. e. the restriction due to the ordinary 256 intensity levels for each colour channel. Most HDR imaging algorithms are based on using multiple exposures of a scene with different settings for each exposure and then using different approaches for storing and displaying this extended image data. However, HDR techniques are not especially aimed at the kind of low-light level data which we target in this work, where the dynamic range in the input data is in the order of 5 to 20 intensity levels (of 255) and the signal-to-noise ratio is extremely low.

Only few published studies explicitly target image enhancement in low light-level video. Two recent examples are the method by Bennett and McMillan (2005) and the technique presented by Lee et al. (2005). Lee et al. (2005) combine very simple operations in a system presumably developed for easy hardware implementation, for example in mobile phone cameras and other compact digital video cameras. According to our tests of this technique (see also Section 7.1) it cannot handle the high levels of noise that we consider in our work.

The approach taken by Bennett and McMillan (2005) for low dynamic range image sequences is more closely connected to our technique. Their *virtual exposures* framework includes the bilateral ASTA-filter (*Adaptive Spatio-Temporal Accumulation*) and a contrast enhancement technique (also called *tone mapping*). The ASTA-filter, which favours relatively more spatial summation and less temporal summation when motion is detected, is in this respect similar to the biological model of Warrant (1999). Since bilateral filters are applied (Tomasi and Manduchi 1998; see above), the filtering is edge-sensitive and the temporal bilateral filter is additionally used for the local motion detection. The filters apply novel dissimilarity measures to deal with the noise sensitivity of the original bilateral filter formulation. The size of the filter applied at each point (i. e. the number of pixels in the local neighbourhood used for the filtering) is determined by an initial calculation of the suitable amount of amplification needed at this point. This amount of amplification is derived by enhancing the contrast in an isotropically smoothed version of the image and recording the amount of amplification used during this procedure. A drawback of the ASTA-fil-

ter is that unlike our technique, it requires global motion detection as a pre-processing step to be able to deal with moving camera sequences. The sequence is then warped to simulate a static camera sequence.

The work presented in the following sections reviews and extends the descriptions of our technique that have been presented previously (Malm and Warrant 2006; Malm et al. 2007).

3. Contrast enhancement

The very first step in our image enhancement approach is to apply an amplifying intensity transformation such that the range of intensities used in the dark areas of the input images is increased while the structure in brighter areas is preserved. We refer to this part of the algorithm as the *contrast enhancement* step. The term *tone mapping* is sometimes also used for this procedure.

In the virtual exposures method of Bennett and McMillan (2005), a contrast enhancement procedure is applied where the intensity transformation is a logarithmic function similar to the one proposed by Drago et al. (2003). The method also contains additional smoothing using spatial and temporal bilateral filters and an attenuation of details, found by the subtraction of a filtered image from a non-filtered image. We instead chose to do all smoothing in the separate noise reduction stage and here concentrate on the contrast enhancement.

The method of Bennett and McMillan relies on several parameters, some of which are used for specifying the bilateral smoothing filters, and some of which are used for changing the influence of two different mapping functions (a function for the large scale data and a function for the attenuation of the details). These parameters have to be set manually and will not adapt when the lighting conditions change in the image sequence. Since we aim for an automatic procedure we instead opted for a modified version of the well-known procedure of histogram equalization (Gonzalez and Woods 1992). Histogram equalization is parameter-free and increases the contrast in an image by finding an intensity transformation that evens out the intensity histogram of the input image as much as possible. However, for many images histogram equalization transforms too much, saturating the brightest intensities so that structure information is lost. We therefore applied contrast-limited histogram equalization as presented by Pizer et al. (1987), but without the tiling that applies different transformations to different areas (tiles) in the image. In contrast-limited histogram equalization, a parameter, the clip-limit β, sets a limit on the derivative of the slope of the transformation function. If the function, found by histogram equalization, exceeds this limit at some point the slope is set to the maximum allowed value and the loss in total increment is counteracted by a uniform increase of the slope over the rest of the function.

4. Noise reduction

After contrast enhancement, the noise in the image sequence will be very apparent, especially in the previously dark areas where there is a low signal-to-noise ratio and intensities have been highly amplified. As mentioned above, we apply adaptive spatio-temporal averaging to reduce this noise while keeping the important edges and structures intact.

We will now review the structure-adaptive anisotropic image filtering method of Yang et al. (1996) and present the modifications and extensions we applied to improve its applicability and to adjust it for our low-light vision objective.

The method computes a new image f_{out} by applying at each spatio-temporal point $\bar{x}_0 = (x_0,y_0,t_0)$ a kernel $k(\bar{x}_0,\bar{x})$ to the original image f_{in} such that:

$$f_{out}(\bar{x}_0) = \frac{1}{\mu(\bar{x}_0)} \iiint_{\Omega} k(\bar{x}_0,\bar{x}) f_{in}(\bar{x}) d\bar{x}, \quad (1)$$

where

$$\mu(\bar{x}_0) = \iiint_{\Omega} k(\bar{x}_0,\bar{x}) d\bar{x} \quad (2)$$

is a normalizing factor. The normalization makes the sum of the kernel elements equal to 1 in all cases, so that the mean image intensity doesn't change. The area Ω over which the integration (or summation in the discrete case) is performed is chosen as a finite neighbourhood centred around $\bar{x}_0$.

Since we want to adapt the filtering to the spatio-temporal intensity structure at each point, in order to reduce blurring over spatial and temporal edges, we calculate a kernel $k(\bar{x}_0,\bar{x})$ individually for each point $\bar{x}_0$. The kernels should be wide in directions of homogeneous intensity and narrow in directions with important structural edges. To find these directions, the intensity structure was analyzed by the so-called *structure tensor* or *second-moment matrix*. This mathematical object has been developed and applied in image analysis in numerous papers by, for example Jähne (1993). The tensor $J_\rho(\bar{x}_0)$ is defined in the following way:

$$J_\rho(\nabla f(\bar{x}_0)) = G_\rho * \left[\nabla f(\bar{x}_0) \nabla f(\bar{x}_0)^T \right], \quad (3)$$

where

$$\nabla f(\bar{x}_0) = \left[\frac{\partial f}{\partial x_0} \quad \frac{\partial f}{\partial y_0} \quad \frac{\partial f}{\partial t_0} \right]^T \quad (4)$$

is the spatio-temporal intensity gradient of f at the point $\bar{x}_0$. G_ρ is the Gaussian kernel function

$$G_\rho = \frac{1}{\mu} e^{-(x^2+y^2+t^2)/2\rho^2}, \quad (5)$$

where μ is the normalizing factor. The notation * in this case means elementwise convolution with the matrix $\nabla f(\bar{x}_0)\nabla f(\bar{x}_0)^T$ in a neighbourhood centered at $\bar{x}_0$. It is this convolution – that gives us smoothing in the direction of gradients – that is the key to the noise insensitivity of this method.

Eigenvalue analysis of J_ρ will now give us the information about the intensity structure that we seek. The eigenvector v_1 corresponding to the largest eigenvalue λ_1 will be approximately parallel to the direction of maximum intensity variation while the other two eigenvectors are orthogonal to this direction. The magnitude of each eigenvalue will be a measure of the amount of intensity variation in the direction of the corresponding eigenvector. For a deeper discussion on eigenvalue analysis of the structure tensor see Haussecker and Spies (1999).

The basic form of the kernel $k(\bar{x}_0,\bar{x})$ constructed at each point $\bar{x}_0$ is a Gaussian function,

$$k(\bar{x}_0,\bar{x}) = e^{-\frac{1}{2} r(\bar{x}_0,\bar{x})}, \quad (6)$$

where

$$r(\bar{x}_0,\bar{x}) = (\bar{x}-\bar{x}_0)^T R \Sigma^2 R^T (\bar{x}-\bar{x}_0) \quad (7)$$

including a rotation matrix R and a scaling matrix Σ. The rotation matrix is constructed from the eigenvectors v_i of J_ρ,

$$R = \left[v_1 \quad v_2 \quad v_3 \right], \quad (8)$$

whereas the scaling matrix has the following form

$$\Sigma = \begin{bmatrix} \frac{1}{\sigma(\lambda_1)} & 0 & 0 \\ 0 & \frac{1}{\sigma(\lambda_2)} & 0 \\ 0 & 0 & \frac{1}{\sigma(\lambda_3)} \end{bmatrix}. \quad (9)$$

The function $\sigma(\lambda_i)$ is a decreasing function of λ that sets the width of the kernel along each eigenvalue direction. The theory of Yang et al. (1996) was mainly developed for 2D images. It calculates measures of corner strength and anisotropism, both involving ratios of the maximum and minimum eigenvalues, at every point $\bar{x}_0$. An extension to the 3D case was discussed. However, we found these two measures to be inadequate for the 3D case since they focus too much on singular corner points in the video input and largely disregard the linear and planar structures that we want to preserve in the spatio-temporal space. For example, a dependence of the kernel width in the temporal direction on the eigenvalues corresponding to the spatial directions, which exists in those measures, is inappropriate in a static background area. Instead, in our own approach an exponential function depends directly on the eigenvalue λ_i in the current eigenvector direction v_i in the following way,

$$\sigma(\lambda_i, \bar{x}_0) = \Delta\sigma e^{-\lambda_i/d} + \sigma_{\min}, \quad (10)$$

where $\Delta\sigma = (\sigma_{max} - \sigma_{min})$, so that σ_i attains its maximum σ_{max} at $\lambda = 0$ and asymptotically approaches its minimum σ_{min} when $\lambda \to \infty$. The parameter d scales the width function along the λ-axis and has to be adjusted to the current noise level. Since the fraction of noise due to the quantum nature of light (photon shot noise) depends on the brightness level, it is signal-dependent and the parameter d should ideally be set locally using a local noise measurement. We are currently developing techniques for automatically setting the parameter d, but in this presentation we will settle for a global setting.

When the widths $\sigma(\lambda_i, \bar{x}_0)$ have been calculated and the kernel subsequently constructed according to equation (6), equation (1) is used to calculate the output intensity f_{out} of the smoothing stage in pixel $\bar{x}_0$.

4.1 Sharpening extension

The structure tensor is surprisingly good at finding the direction of motion even in very noisy input sequences. However, in some cases the elongation of the constructed summation kernels can be slightly misaligned relative to the motion direction. In these cases the contours of the moving objects can be somewhat blurred. As an alternative to our standard adaptive smoothing approach we propose the addition of so-called *high-boost filtering* to sharpen up these contours. A spatio-temporal high-boost filter is defined as a 3 × 3 × 3 tensor with the value –1 in all elements except the center element that is ascribed the value 27A-1, where A is a constant (Gonzalez and Woods 1992). We have attained the best results using A = 1.2. However, constructing the filter in this way and applying it after the smoothing process takes additional time. It is more time efficient to incorporate the high-boost filter directly into the smoothing kernel. We have done this in the following manner.

The smoothed image response f_{out} is given by convolving the image f_{in} with the anisotropic kernel k according to equation (1). In the high-boost version we use a convolution with a high-boost filter H. So the new kernel $\hat{k}$ is given by

$$\hat{k} = H * k, \quad (11)$$

where * now means ordinary convolution. We have used the following continuous approximation to the discrete high-boost filter:

$$H = e^{-(x^2+y^2+t^2)/2\omega^2} \\ (1-\frac{\alpha^2}{2}(x^2+y^2+t^2)). \quad (12)$$

Here, α and ω are two constants that are chosen so that this analytic function resembles the discrete version of the filter as much as possible. In practice, we have chosen α = 1.55 and ω = 0.515, or values close to these.

The convolution in equation (11) can be performed analytically so that we get a closed form for the combined smoothing and high-boost filter:

$$\hat{k}(\bar{x}) = e^{-\frac{1}{2}h_1(\bar{x})} h_2(\bar{x}), \quad (13)$$

with

$$h_1(\bar{x}) = (\frac{x^2}{\sigma_1^2+\omega^2} + \frac{y^2}{\sigma_2^2+\omega^2} + \frac{t^2}{\sigma_3^2+\omega^2}), \quad (14)$$

and

$$h_2(\bar{x}) = 1 - \frac{3\alpha^2\omega^2}{2} - \frac{\alpha^2\omega^4}{2} \\ (\frac{x^2-1}{\sigma_1^2+\omega^2} + \frac{y^2-1}{\sigma_2^2+\omega^2} + \frac{t^2-1}{\sigma_3^2+\omega^2}). \quad (15)$$

This makes it possible to perform both the smoothing and the sharpening with one summation just by changing the kernel from k to $\hat{k}$ in equation (1).

Applying this new kernel in the spatio-temporal summation step clearly improves the output. The images look crisp and sharp and free of noise. The simultaneous combination of smoothing and sharpening also effectively minimizes spurious detail, as well as ringing effects near the most distinct edges, which otherwise often appear after an image sharpening procedure.

4.2 Stretching

After the noise reduction step, the output sometimes appears to have lost some of the contrast previously gained in the contrast enhancement step. The reason for this is that the adaptive smoothing has a tendency to squeeze the intensities into the central part of the dynamic range, so that the peripheral intensity values are scarcely used. This effect can be alleviated by first setting the pixels with the lowest intensity values to zero. Let us assume that the number of pixels with intensity values of up to 50 (of 256) only corresponds to around 0.1 % of the total number of pixels. Choosing to set those pixels to zero would not result in a significant loss in information. After this we then stretch the colour map uniformly. Since we target very dark sequences, it is usually not a good idea to apply the same procedure for the brightest part of the intensity range. This is because there are often small bright details, such as a single lamp, that clearly change appearance from the input to the output after such an operation.

5. Consideration for colour

The discussion so far has dealt with intensity images, but the algorithm developed here also works well for data on coloured images. It can restore the original chromaticity values along with the luminance. We will here discuss some special aspects of the algorithm in regard to processing colour images.

When applying the algorithm to RGB colour image data one could envision a procedure where the colour data in the images are first transformed to another colour

space including an intensity channel, for instance the HSV (Hue, Saturation, Value) colour space (Gonzalez and Woods 1992). The algorithm could then be applied unaltered to the intensity channel, while smoothing of the other two channels could either be performed with the same kernel as in the intensity channel or by isotropic smoothing. The HSV image would then be transformed back to the RGB colour space.

However, in very dark colour video sequences the noise levels in the different input channels often differ significantly from each other. For example, the blue channel often has a relatively high noise level. It is therefore essential that there is a possibility to adapt the algorithm to this difference. To this end, we chose to calculate the structure tensor J_ρ, and its eigenvectors and eigenvalues in the intensity channel, which we simply define as the mean of the three colour channels. The widths of the kernels are then adjusted separately for each colour channel by using a different value of the scaling parameter d for each channel. Changing d affects the amount of smoothing, but the mean intensity in the local neighbourhood stays the same. This improves the output with colours closer to the true chromaticity values and with less false colour fluctuations than those resulting from the above mentioned HSV approach.

When acquiring raw image data from a CCD or CMOS sensor, the pixels are usually arranged according to the so-called Bayer pattern. As shown by Hirakawa and Parks (2006), the interpolation from the Bayer pattern to three separate channels, the so-called *demosaicing*, can be efficiently applied simultaneously to the de-noising of the image data. We apply this approach here for each channel. For pixels where intensity data are not available, we set the coefficients in the summation kernel $k(\overline{x}_0,\overline{x})$ to zero and then normalize the kernel. A smoothed output is then calculated for both the noisy input pixels and the pixels where data are missing.

6. Implementation on a graphics processing unit

Filtering using the summation kernels $k(\overline{x}_0,\overline{x})$ is inherently amenable to parallel processing, and for this the *graphics processing units* (GPUs) of modern graphics cards are well suited. We have implemented the whole adaptive enhancement methodology as a combined CPU/GPU algorithm.

All image pre- and post-processing is performed on the *central processing unit* (CPU). This includes image input/output and the contrast enhancement step. The histogram equalization, which implements the contrast enhancement, requires summation over all pixels. This computation is not easily adapted to a GPU, as the summation would have to be done in multiple passes. However, as these steps constitute only a small amount of the execution time, a simpler CPU implementation suffices here.

The most time-expensive steps of the complete algorithm are the calculation of the structure tensor, including the gradient calculation, the element-wise smoothing, and the actual filtering, or summation, all of which we perform entirely on the GPU. To calculate the gradients we first upload n frames of the input sequence to the graphics card as floating-point 2D textures, and then perform spatial gradient computations on each of these. Next we take temporal differences between successive frames and use the resulting gradient to compute the structure tensor for each pixel in each frame. Using the separability of the Gaussian kernel, we then compute the isotropically smoothed tensor for each frame. Typically

these filtering steps are performed using filter sizes of up to 7 × 7 × 7 pixels. For the normalization of the filter, the alpha channel is used to store the sum of the filter weights.

The smoothing kernels are then computed in a fragment program on the GPU. This involves finding the eigenvalues and eigenvectors of the structure tensor, which can be done efficiently in a single pass on modern GPUs. The kernel coefficients are temporarily stored as textures and the final spatio-temporal summation is performed (similarly to the isotropic pre-pass) in multiple 2D passes. As the filter kernels are unique for each pixel, the filter weights are re-computed on the fly during the filtering process.

In summary, by exploiting the massively parallel architecture of modern GPUs, we obtain real-time frame rates on a single nVidia GeForce GTX 260 graphics card. Using 7 × 7 × 7 pixel neighbourhoods for the spatio-temporal smoothing kernels, and a resolution of 300 × 300 pixels, we achieve a frame rate of approximately 30 frames per second. The ratio of times spent on computations on the CPU, on the structure tensor computations and on the adaptive spatio-temporal smoothing are approximately 1 : 2 : 6.

7. Results

We have tested our complete enhancement method on a variety of gray-scale and colour image sequences, differing in resolution and light level and obtained by both consumer grade and machine vision cameras. The tested input sequences include movies obtained by stationary cameras as well as by moving cameras. Overall, the method gives strikingly clean and bright output, considering the bad quality of the input data.

Figure 1 presents one frame of a colour sequence obtained with a Sony DCR-HC23

Fig. 1 Upper left: One frame of the original sequence. Upper right: Contrast enhanced version. Lower left: After noise reduction. Lower right: After intensity stretching

hand-held consumer camcorder. The upper left image is the original dark sequence. The resolution of this clip is 360 × 288. On the upper right, the contrast-enhanced image is seen after the application of the contrast-limited histogram equalization described in Section 3. The image in the lower left corner is taken from the sequence following the adaptive smoothing as described in Section 4. The smoothing decreases the contrast somewhat, but this degradation is automatically alleviated by the stretching procedure discussed in Section 4.2. The improvement following this last step is shown in the lower right corner of Fig. 1.

Figure 2 shows the algorithm at work on a 640 × 480 colour sequence taken by a moving camera mounted on a model railway train. The camera used is an AVT Marlin machine vision camera. The input to the algorithm is the raw sensor data and therefore the demosaicing of the Bayer pattern is done simultaneously to the de-noising. As the camera moves, a bird spider is moving from left to right in the scene. The algorithm has no problem in smoothing the intensity in the best possible way for these different motions, and the image of the spider is well preserved. The output shown is the result of the algorithm with the sharpening extension. The method preserves the sharp details of the buildings while reducing the noise effectively.

7.1 Comparison to other methods

We compared our method mainly with two other methods that also deal with low-light

Fig. 2 Upper left: One frame from the original raw input data. Upper right: Contrast enhanced version with standard demosaicing. Lower panel: after noise reduction, including demosaicing and sharpening

Fig. 3 From left to right: **1** A frame from a dark input sequence; **2** after contrast enhancement according to Section 3; **3** output of our enhancement method, without sharpening; **4** output of the method by Lee et. al. (2005); and **5** output after the application of the method by Bennett and McMillan (2005)

level conditions. Firstly, we implemented the enhancement method presented by Lee et al. (2005). In this method, Poisson noise is removed in stationary areas by calculation of the median in a small spatial neighbourhood. So-called false colour noise is removed by either choosing the intensity value of the current pixel, or the value of the preceding corresponding pixel, depending on a simple measure of the variance in a neighbourhood of the pixel. For the very low light levels that we are dealing with here, and the high noise levels that this implies, the output of the method of Lee et al. was visually very disappointing. This does not surprise us since the method is aimed at efficient real-time implementation on hardware platforms and only uses 3 × 3 pixel neighbourhoods and two succeeding images for the calculations. The calculations of the medians in these small neighbourhoods leave heavy flickering in the output.

The competing method giving the best output was the "virtual exposures" method proposed by Bennett and McMillan (2005). This method was briefly described in Sections 2 and 3. When a sequence of images is obtained by a stationary camera, the virtual exposure method mostly results in temporal filtering. This removes a large part of the heavy noise in the input sequence. However, the output isn't as spatially smooth as the output resulting from the application of the method presented in this paper. When the camera is moving, an external motion estimation is needed. Moreover, the contrast enhancement achieved by the tone mapping is disappointing for the cases we tested. We suspect that this is due to the fact that Bennett's and McMillan's tone mapping approach uses separate processing of the low and high frequency parts of the input data which isn't optimal for the extreme levels of noise that we are dealing with here.

Figure 3 shows some results of the processing of a very dark grey-level image sequence obtained with a Sony XC-50 machine vision camera. From left to right, Fig. 3 shows a frame from the dark input sequence, and then the same frame after 1) contrast enhancement according to Section 3; 2) application of our enhancement method, without sharpening; 3) application of the method by Lee et al. (2005); and 4) application of the method of Bennett and McMillan (2005).

Conclusions

We have here presented a methodology for adaptive enhancement and noise reduction for very dark image sequences with very low dynamic range. The method has been inspired by spatio-temporal summation principles used in the visual systems of nocturnal animals and in general follows the same principle process-

ing steps that are used in the animal visual pathway. The approach is very general and adapts to the spatio-temporal intensity structure in order to prevent motion blur and smoothing across important structural edges. The method also includes a sharpening feature that prevents the most important object contours from being over-smoothed.

Most parameters can be generally set for a very large group of input sequences. These parameters include: the clip-limit in the contrast-limited histogram equalization, the maximum and minimum widths of the filtering kernels and the width of the isotropic smoothing of the structure tensor and smoothing in the gradient calculations. However, the scaling parameter for the width function has to be adjusted to the noise level in the current sequence. The best approach when applying the method to colour images has been discussed, and this includes demosaicing from the Bayer pattern in raw input colour data simultaneously to the noise reduction. We have implemented the method using a GPU and achieved real-time performance. For very noisy video input data, which is the result of filming in very low light levels, the method presented here outperforms (in terms of output quality) all competing methods that we have encountered at the time of the writing of this chapter.

Further work on this technology aims at the reduction of the number of computations per frame, so that real-time performance on high resolution images becomes possible. This might be done by calculating the adaptive smoothing kernels only within a sparse grid and then interpolating these calculated kernels for the remaining pixels. Another important aspect to look into is an automatic local setting of the scaling parameter *d* for the width function. This parameter should be related to a local noise estimation around every processed pixel in order to optimize the amount of smoothing.

References

Alter F, Matsushita Y, Tang X (2006) An intensity similarity measure in low-light conditions. Proc Europ Conf Computer Vision, 267–280

Barash D, Comaniciu D (2004) A common framework for nonlinear diffusion, adaptive smoothing, bilateral filtering and mean shift. Image and Vision Computing 22: 73–81

Bayer B (1976) Color imaging array. US patent 397 165

Bennett E, McMillan L (2005) Video enhancement using per-pixel virtual exposures. Proc SIGGRAPH, Los Angeles, CA: 845–852

Boo K, Bose B(1998) A motion-compensated spatio-temporal filter for image sequences with signal-dependent noise. IEEE Trans Circuits Syst Video Tech 8(3): 287–298

Boyce J (1992) Noise reduction of image sequences using adaptive motion compensated frame averaging. Proc IEEE Int Conf Acoust Speech Signal Proc 3: 461–464

Bruhn A, Wieckert J, Schnörr C (2005) Lucas Kanade meets Horn/Schunk: Combining local and global optic flow methods. Int J Computer Vision 61(3): 211–231

Drago F, Myszkowski K, Annen T, Chiba N (2003) Adaptive logarithmic mapping for displaying high contrast scenes. Proc EUROGRAPHICS 22: 419–426

Dugad R, Ahuja N (1999) Video denoising by combining Kalman and Wiener estimates. Proc Int Conf Image Proc 4, Kobe, Japan: 152–156

Frederiksen R, Wcislo WT, Warrant EJ (2008) Visual reliability and information rate in the retina of a nocturnal bee. Current Biol 18: 349–353

Gonzalez R, Woods R (1992) Digital image processing. Addison-Wesley, Reading MA

Granlund G, Knutsson H (1995) Signal processing for computer vision. Kluwer, Dordrecht

Hartline HK, Ratliff F (1957) Inhibitory interaction of receptor units in the eye of *Limulus*. J Gen Physiol 40: 357–376

Hateren JH van (1992) Real and optimal neural images in early vision. Nature 360: 68–70

Hateren JH van (1993) Spatiotemporal contrast sensitivity of early vision. Vision Res 33: 257–267

Haussecker H, Spies H (1999) Handbook of computer vision and applications. Academic Press, New York London: 125–151

Hirakawa K, Parks T (2006) Joint demosaicing and denoising. IEEE Trans Image Processing 15(8): 2146–2157

Horn B (1986) Robot vision. MIT Press, Cambridge MA

Jähne B (1993) Spatio-temporal image processing. Springer, Berlin

Kalivas D, Sawchuk A (1990) Motion compensated

enhancement of noisy image sequences. Proc IEEE Int Conf Acoust Speech Signal Proc, Alburquerque, NM: 2121–2124

Kim J, Woods J (1997) Spatio-temporal adaptive 3-d Kalman filter for video. IEEE Tran Image Proc 6(3): 414–424

Laughlin SB, Weckström M (1993) Fast and slow photoreceptors: A comparative study of the functional diversity of coding and conductances in the Diptera. J Comp Physiol A 172: 593–609

Lee S, Kang M (1998) Spatio-temporal video filtering algorithm based on 3-d anisotropic diffusion equation. Proc Int Conf Image Processing 2: 447–450

Lee SW, Maik V, Jang J, Shin J, Paik J (2005) Noise-adaptive spatio-temporal filter for real-time noise removal in low light level images. IEEE Tran. Consumer Electronics 51(2): 648–653

Malm H, Warrant E (2006) Motion dependent spatiotemporal smoothing for noise reduction in very dim light image sequences. Proc 18th Int Conf Pattern Recognition, Hong Kong: 954–959

Malm H, Oskarsson, M, Warrant E, Clarberg P, Hasselgren J, Lejdfors C (2007) Adaptive enhancement and noise reduction in very low light-level video. Proc 11th Int Conf Computer Vision, Rio de Janeiro, 1–8

Miyata K, Taguchi A (2002) Spatio-temporal seperable data-dependent weighted average filtering for restoration of the image sequences. Proc IEEE Int Conf Acoust Speech Signal Proc, Alburquerque, NM: 3696–3699

Özkan M, Sezan M, Tekalp A (1993) Adaptive motion-compensated filtering of noisy image sequences. IEEE Tran Circuits Sys Video Tech 3(4): 277–290

Perona P, Malik J (1990) Scale-space and edge detection using anisotropic diffusion. IEEE Trans on Patt. Analysis and Machine Intelligence 12(7): 629–639

Pizer S, Amburn E, Austin J, Cromartie R, Geselowitz A, Geer T, teer Haar Romeny B, Zimmerman J, Zuiderveld K (1987) Adaptive histogram equalization and its variations. Computer Vision, Graphics and Image Processing 39: 355–368

Reibel Y, Jung M, Bouhifd M, Cunin B, Draman C (2003) CCD or CMOS camera noise characteristics. Proc Euro Physical J of Applied Physics: 75–80

Sezan M, Özkan M, Fogel S (1991) Temporally adaptive filtering of noisy image sequences. Proc IEEE Int Conf Acoust Speech Signal Proc 4: 2429–2432

Stockham T (1972) Image processing in the context of a visual model. Proc of the IEEE 60(7): 828–842

Tomasi C, Manduchi R (1998) Bilateral filtering for gray and color images. Proc 6th Int Conf Computer Vision: 839–846

Uttenweiler D, Weber C., Jähne B, Fink R, Scharr H (2003) Spatiotemporal anisotropic diffusion filtering to improve signal-to-noise ratios and object restoration in fluorescence microscopic image sequences. J Biomedl Optics 8(1): 40–47

Warrant EJ (1999) Seeing better at night: life style, eye design and the optimum strategy of spatial and temporal summation. Vision Res 39: 1611–1630

Warrant EJ (2004) Vision in the dimmest habitats on earth. J Comp Physiol A 190: 765–789

Warrant, EJ (2008) Nocturnal vision. In: Albright T, Masland RH (eds) The senses: A comprehensive reference (Vol. 2: Vision II). Acad Press, Oxford: 53–86

Weickert J (1998) Anisotropic diffusion in image processing. Teubner-Verlag, Stuttgart

Weickert J (1999) Coherence-enhancing diffusion filtering. Int J Computer Vision 31(2/3): 111–127

Yang G, Burger P, Firmin D, Underwood S (1996) Structure adaptive anisotropic image filtering. Image and Vision Computing 14: 135–145

Event-based silicon retinas and cochleas

6

Tobi Delbruck, Shih-Chii Liu

Contents

Abstract

This chapter reviews neuromorphic silicon retinas and cochleas that are based on the structure and operation of their biological counterparts. These devices are built using conventional chip fabrication technologies, using transistor circuits that emulate neural computations from biology. In first generation sensors, the analog outputs of every cell were read out serially at fixed sample rates. The new generation of sensors reports only the outputs of active cells through digital events (spikes) that are communicated asynchronously. Such sensors respond more quickly with reduced power consumption. Their digital "address-event" outputs rapidly convey precise timing information about the scene that is only attained from conventional sensors if they are continuously sampled at high rates. The sparseness, low latency, and spatio-temporal structure of this new form of sensor output data can benefit subsequent post-processing algorithms. Tradeoffs in the design of neuromorphic visual and auditory sensors are discussed. Examples are given of vision algorithms that process the address-events, using their spatio-temporal coherence, for low-level feature extraction and for object tracking.

Tobi Delbruck
University of Zürich and ETH Zürich
Institute of Neuroinformatics
Winterthurerstr. 190, 8057 Zürich, Switzerland
e-mail: tobi@ini.phys.ethz.ch

1. Introduction

Biology provides examples of efficient machines which greatly outperform conventional technology. Electronic neuromorphic systems aim to capture the same efficient style of computation by emulating the struc-

ture and function of biological counterparts (Mead 1990; Douglas et al. 1995). This chapter discusses recent progress in realizing neuromorphic sensory systems which mimic the biological retina and cochlea, and subsequent sensor processing. The new generation of electronic sensors communicates their activity in much the same way as nervous systems do, by asynchronously outputting streams to digital events representing spikes. Because sensory computation is driven by these events, which originate from the sensory input, such systems are called event-driven systems. The progress in two specific neuromorphic sensors, the retina and the cochlea, is described in this chapter along with examples of how these sensors can help solve challenging problems in vision and audition.

How can asynchronous spike-based communication in neuromorphic systems be implemented? In contrast to biology which uses a slow ionic medium to carry spikes on dedicated axonal connections, event-based systems use the high speed of electrons to transmit source spike addresses from multiple sources (e. g. retina pixels) on a single device (e. g. a silicon retina) onto shared digital buses. Receivers decode these addresses for further use. The name for this asynchronous transmission scheme is the address-event representation, or **AER**. Using AER is a natural match to sensors that reduce redundancy in the input, thus transmitting only sparse useful information. Using AER, outputs can be transmitted with short response latencies and at a low data rate. Both characteristics are useful for practical applications, and they also differentiate AER sensors from conventional sensors that sample redundant information at constant rate.

This chapter begins with the history of silicon retina designs and then describes in detail the design and operation of an example AER retina. Section 3 discusses the history and progress of silicon cochleas. Section 4 discusses how the address events from the retina are processed in vision applications, with a focus on the use of the event timing; and Section 5 summarizes the chapter.

2. Conventional cameras and silicon retinas

The continued shrinking of silicon transistors has enabled rapid development of conventional frame-based electronic camera technology (Fossum 1997). With pixel sizes barely over a micron nowadays, a naïve observer might think that industry can approach a “perfect” image sensor with infinite resolution, infinite dynamic range, infinite frame rate, zero pixel size, and zero power consumption. Of course the output from such an ideal sensor would be impossible to process. As neuromorphic engineers, we observe that biological vision relies on a sensor (the retina) that does local gain control and massive amounts of computation to produce at its output (the optic nerve), an asynchronous stream of digital data (spikes) that represents only relevant visual information. A vision sensor that emulates its biological counterpart could help solve vision problems that seem easy for biology but are difficult using present machine vision techniques.

2.1 History of retina designs

The first integrated silicon retina by Mahowald and Mead (1991) consists of a 3-layered model of the outer retina architecture that included the photoreceptors, bipolar cells, and horizontal cells. The spiking ganglion cells and AER interface were incorpor-

ated in a later design. The key biological features emulated in this retina include the logarithmic, adaptive, phototransduction stage, the horizontal cell spatio-temporal smoothing network, and the rectification of the retina output into complementary ON and OFF channels. This retina was also part of the first AER vision system built by Mahowald and colleagues to compute stereopsis (Mahowald 1992, 1994).

However, this and subsequent AER systems in the following decade were barely useable. Mahowald's retina for instance, would respond reliably only to a flashing LED held up in front of it. Following this retina, Boahen's group (Zaghloul and Boahen 2004) built a more biologically realistic version with four cell types emulating the ON and OFF cell types for the sustained and transient pathways, but it suffered from large pixel response variability. The vision sensor by Ruedi et al. (2003) outputs simultaneous spatial contrast and local orientation using an event-based readout in which the pixel event timing within a frame encodes local contrast information. This design was the first that showed it was possible to overcome the mismatch, poor dynamic range, and otherwise low performance that plagued prior retinas.

Other recent event-based retinas that aim to reduce image redundancy include spatial contrast AER retinas (Barbaro et al. 2002; Ruedi et al. 2009; Lenoro-Bardallo et al. 2010), temporal intensity change retinas (Mallik et al. 2005), temporal-contrast retinas (Kramer 2002; Lichtsteiner et al. 2004), the dynamic vision sensor silicon retina described in Section 2.3 (Lichtsteiner et al. 2008), a prototype temporal-contrast infrared bolometer (Posch et al. 2009), preliminary color retina pixel designs (Fasnacht and Delbruck 2007; Berner et al. 2008; Olsson and Hafliger 2009), and a recent design which outputs both temporal contrast and gray level information (Posch et al. 2010). These designs need to cope with the fact that fabricated transistor characteristics are quite variable.

2.2 Achieving precision sensing with imprecise transistors

One of the biggest impediments in bringing the neuromorphic approach to commercially successful products is the problem of transistor mismatch. Transistor mismatch refers to built-in variability in transistor operating characteristics, and leads to the static salt and pepper noise in the output of initial retinas which made the image barely recognizable. Transistor mismatch is inherent to the way in which transistors are fabricated (Pelgrom et al. 1998). It is also one of the two major reasons why the silicon industry revolves around Boolean logic circuits, which reduce the impact of transistor mismatch by restoring the circuit output into one of the two power supply levels of a circuit block. (The other reason is the availability of tools for designing complex synchronous logic circuits.) Because matching precision scales with transistor dimension, precise analog design requires an increase of transistor area devoted to analog processing as process technology scales down. These considerations suggest that analog signals from sensors should be converted quickly to a digital representation so that the subsequent computational precision can be precisely controlled. However, at lower precision, analog computations take less area and consume less power than equivalent digital representations and computations (Sarpeshkar 1998). Analog designers have increasingly focused their efforts on embedding circuit mechanisms that decrease the effect of device imprecision on sensor performance. These mechanisms include dynamic adaptation and gain control that optimally utilizes the internal signal

range, digital calibration, and software post-processing. The most successful neuromorphic sensors have paid particular attention to improvements in their response uniformity. The silicon retina described next uses a combination of techniques to improve its sensitivity.

2.3 The dynamic vision sensor (DVS) silicon retina

The silicon retina discussed in this section is called the dynamic vision sensor or **DVS**, because it emulates retinal processing concerned with dynamic information. The DVS is used as an example of a silicon retina where organizing principles are successfully transferred from biology into silicon. The DVS forms an abstraction of the retina's transient pathway (Fig. 1a), which is blind to static input, monochromatic, and relatively fast (Rodieck 1998). The transient pathway can be contrasted with the sustained pathway, which has properties more like conventional cameras.

The DVS models a simplified 3-layer retina (Fig. 1b) consisting of photoreceptors, bipolar cells, and ganglion cells. In the simplified biological retina in Fig. 1b, the photoreceptors transduce from the input stimulus light to electrochemical signals. Bipolar cells compare the photoreceptor outputs with spatiotemporal lowpass averages computed by the horizontal cell spreading network and rectifiy the differences into ON and OFF activity. ON and OFF ganglion cells generate action potentials (spikes) from the bipolar cell outputs. The DVS simplifies this classical scheme even further by omitting the horizontal spatial averaging computed by the horizontal cells. Each DVS pixel (Fig. 1c) consists of 3 parts: a logarithmic photoreceptor, a temporal-change "differencer" amplifier (bipolar cells), and 2 decision units (ganglion cells). The pixel output is an asynchronous stream of ON and OFF events that signal log intensity changes, which usually correspond with scene reflectance changes. The generation of these events follows the process illustrated in Fig. 1d: The continuous-time photoreceptor output encodes intensity logarithmically as a voltage ($\log I$). The differencer amplifies the change in this output from a memorized value to produce $A \cdot d(\log I) = A \cdot dI/I$. Ganglion cell comparators detect a change in log intensity which exceeds a threshold and emit an ON or OFF event. The emission of an event also causes a new $\log I$ value to be memorized across the capacitor. The ON and OFF thresholds are typically set to about 10 % contrast. AER communication circuits along the periphery of the chip (Fig. 1e) transmit the address events off-chip with latencies of less than a microsecond.

What is the meaning of each DVS event? Each event represents a quantized change in log intensity. Light falling on the retina is a product of the scene illumination and reflectance. In the usual condition where scene illumination is constant over time, the static logarithmic response of the DVS photoreceptor turns this product into a sum. The differencing operation that subsequently produces events thereby encodes scene reflectance changes, which typically represent moving objects. Because this computation is based on a compressive logarithmic transformation in the photoreceptor circuit, the pixel can operate over a wide dynamic range of background intensity (120 dB as compared with 60 dB for a high quality conventional image sensor). This wide dynamic range capability means that the sensor can be used under uncontrolled natural lighting conditions that are typified by wide variations in scene illumination. The events are transmitted off chip in a fraction of a microsecond, so the timing of the pixel event is preserved, leading to a typical effective frame rate of several kHz. The data

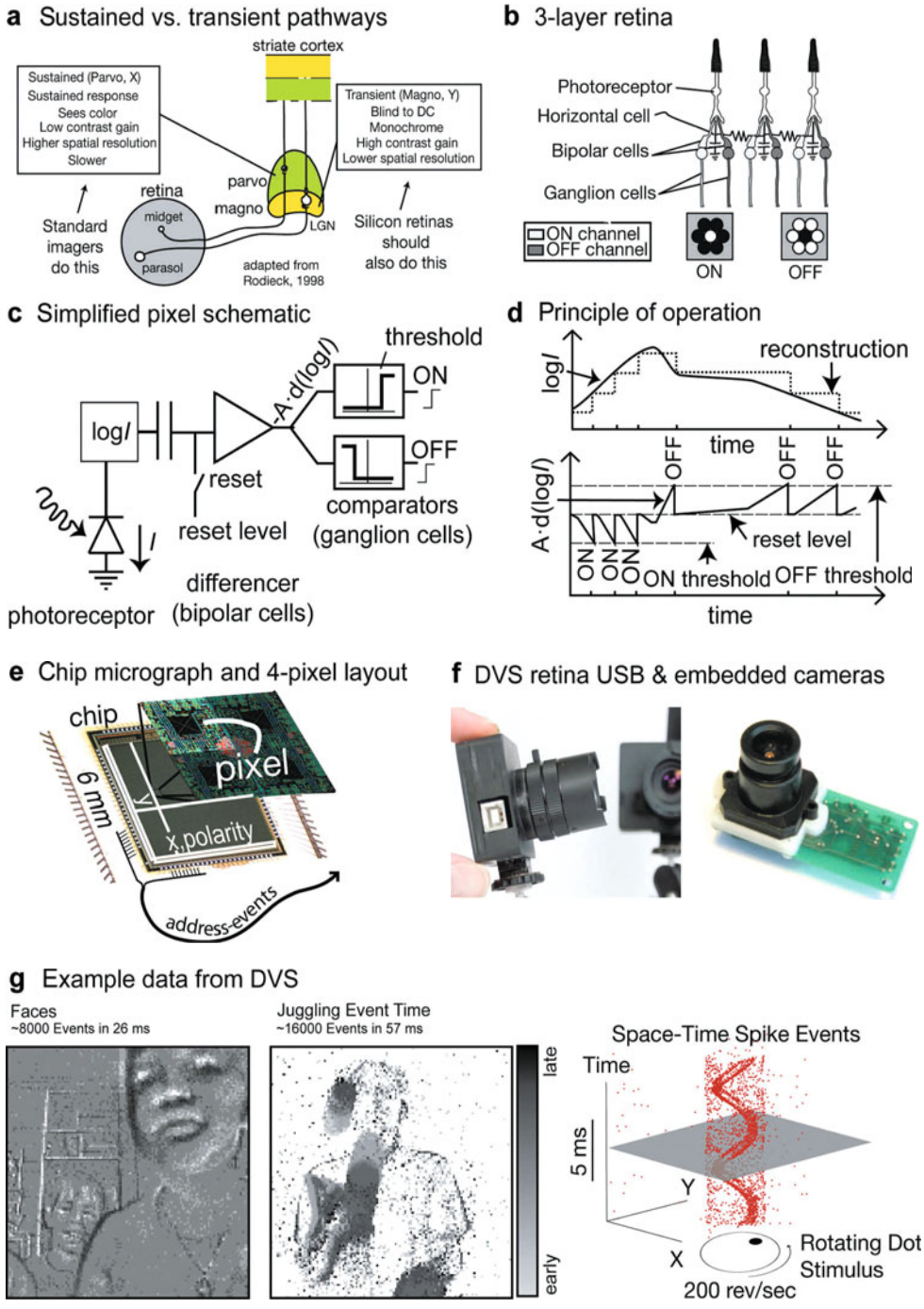


Fig. 1 Dynamic vision sensor (DVS) silicon retina. **a** Sustained and transient pathways from the retina through the lateral geniculate nucleus (LGN) to the striate cortex; **b** 3-layer retina; **c** DVS pixel schematic; **d** principle of operation; **e** DVS die micrograph and mirror symmetric arrangement of 4 pixels; **f** DVS retina USB2.0 camera and embedded microcontroller camera; **g** example DVS output data. "Faces" shows events rendered as a 2 d histogram over a 26 ms time slice. "Juggling Event Time" shows events from a juggler rendered so that lighter events occurred later in the 57 ms time slice. "Space-Time Spike Events" shows events from a spinning dot rendered in 3 d space-time. Adapted from Liu and Delbruck (2010)

rate is often a factor of 100 times lower than that of a frame-based image sensor with an equivalent time resolution. The pixel design provides enough response uniformity such that the pixel contrast threshold can be set to about 10 % contrast, allowing the device to sense real-world contrast signals.

The DVS chip can be connected to a variety of other devices (Fig. 1f). It is only necessary that the receiving device (e. g. custom multi-neuron chips or microcontroller) use the digital AER protocol. The chip has been integrated into a camera unit which interfaces to a computer via a standard USB2.0 interface that delivers time-stamped address-events to a host PC where post-processing of the output events allows us to easily explore event-driven algorithms which we can test in real-world scenarios.

The DVS output data (Fig. 1g) shows interesting characteristics: The "Faces" image shows the 2d histogram of the ON and OFF events over a time slice of 26 ms. The gray level of the "Juggling event time" image represents the event time relative to the start of the time window. The "Space-Time Spike Events" image exposes the spatiotemporal structure of a moving object, in this case a spinning dot, by displaying the events in space-time coordinates.

The DVS retina model is only a loose abstraction of the transient pathway. Even though this simplified silicon retina lacks many classes of retinal cells, it is proving useful for various practical applications as will be discussed in Section 4.

3. Conventional audition and silicon cochleas

Historically, incorporating biological knowledge into machine audition has lead to improved systems. Biologically inspired methods improve the performance of speech recognition systems (Gold and Morgan 2000). However the input stage of these systems is different for that of the biological cochlea. In machines, the acoustic signal is processed by extracting frequency components of the sound within a time window (20 ms for example) using a Discrete Fourier Transform (DFT). The power spectrum of the DFT is transformed logarithmically and again processed through a filter bank which has a log distribution similar to that measured in human cochleas. Systems based on the above methods are dominant because of the efficiency of the Fast Fourier Transform (FFT), and digital signal processing (DSP) chips that support the type of mathematical computation needed for FFTs. Contemporary hearing aids, for example, implement an entire signal processing chain based on a set of customized mostly-digital signal processing chips from microphone to speaker on a power budget of about 1 mW.

However, artificial audition systems still lack the performance of biological systems, for example, in tasks such as understanding speech in noisy backgrounds. Researchers look at replacing the input stage with a continuous-time cochlea processing stage and are also starting to explore the use of spikes for representing signals in speech processing systems, utilizing what is known about speech features in various brain regions (Liu et al. 2010b).

While silicon cochlea designs have been evolving over the last 2 decades; recent cochlea implementations also produce asynchronous spike outputs similar to that of biological cochleas. The analog timing of the silicon cochlea spikes carries precise timing information about the sound, similar to spikes from the spiral ganglion cell outputs of the biological cochlea. Developing methods to process these spikes can guide us on how information is processed from the co-

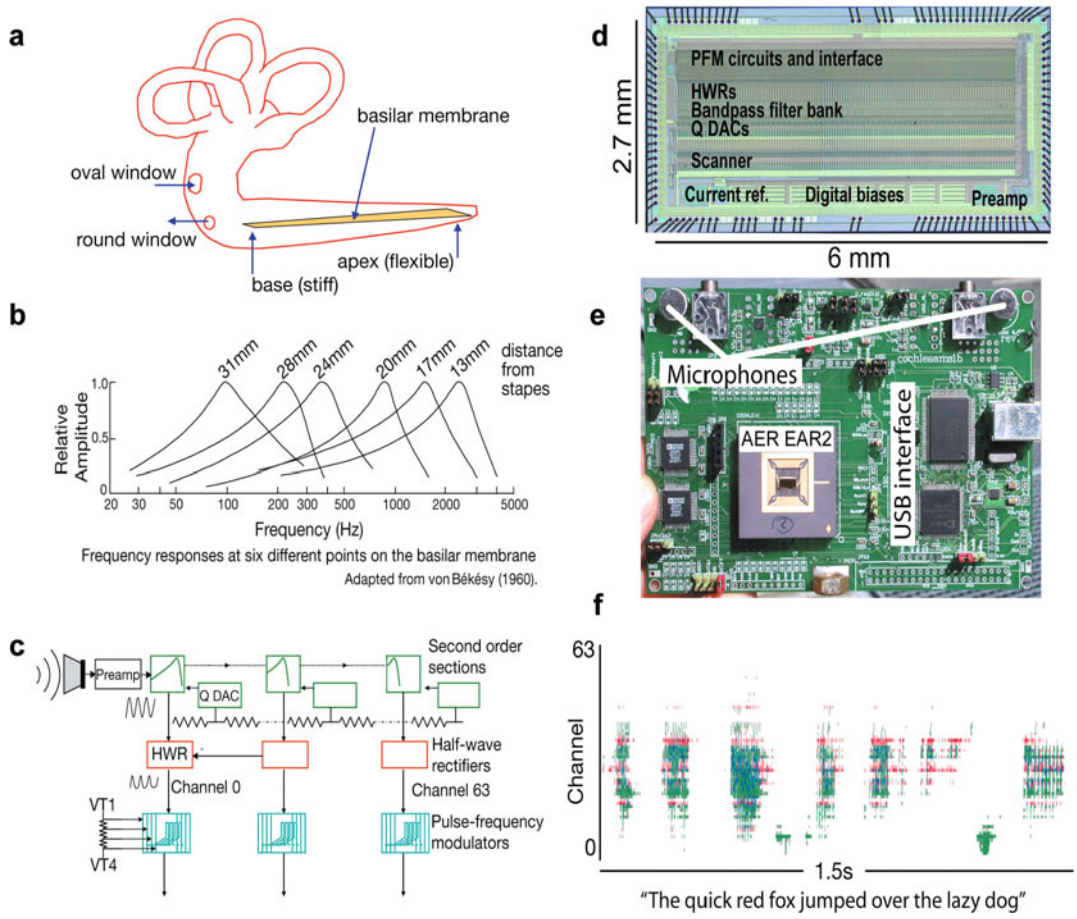


Fig. 2 The AER-EAR2 silicon cochlea. **a** cochlea and its basilar membrane (BM); **b** frequency response of the BM; **c** cascaded second-order-section (SOS) stages of the AER-EAR2 silicon cochlea; **d** integrated binaural 64 channel spike-based cochlea; **e** USB board implementation; **f** response to speech "The quick red fox jumped over the lazy dog". Each dot is an event; lower channel numbers are tuned to higher characteristic frequencies. Adapted from Liu and Delbruck (2010)

chlea to the cochlear nucleus, the midbrain, and other brain structures in the auditory cortex, and might offer insights into how humans are able to process speech and other sounds effectively in adverse environments such as hearing through noise, reverberation and interference from other speakers.

3.1 History of cochlea designs

Silicon cochlea models often implement the basilar membrane (BM) architecture as a cascaded set of second-order-section (SOS) bandpass filters of Lyon and Mead (1988) (see Fig. 2a). Depending on the input frequencies, a frequency component of the pressure wave generated in the fluid leads to a maximum displacement at some place along the basilar membrane (Fig. 2b) depending on its characteristic frequency (CF). The basilar membrane structure has a characteristic frequency (CF) which decreases exponentially from the base to the apex. The cascaded 1-d cochlea captures the sharp roll-off in frequency response of the basilar membrane but has limitations; a failure in one of the stages will lead to a failure in sub-

sequent second-order-section (SOS) stages. Also, the accumulation of noise along the cascaded stages presents a problem. But because the noise outside of the pass band of the filters will be removed as one proceeds along the cascade, the noise accumulation quickly saturates to a level that is tolerable (Sarpeshkar and Lyon 1998). A large number of stages within a frequency range also lead to a larger delay through the cascade than observed in the real cochlea.

Because of these drawbacks, a number of labs have attempted to construct BM models based on a parallel (rather than cascaded) 1-d architecture, or a 2-d architecture that includes the role of the fluid coupling between the stages (Watts et al. 1992; Fragniere 2005; Wen and Boahen 2006; Hamilton et al. 2008). But these approaches have their own drawbacks, such as increased chip area, and sensitivity to mismatches between stages, which cause destructive interference and greatly reduce the theoretically-achievable gain. A recent hot topic in cochlear designs concerns the implementation of local gain control as enabled by the outer hair cells. Developers are considering ways of measuring of the BM signal amplitudes and using these measurements to tune the sharpnesses of the resonances of the stages (Sarpeshkar and Lyon 1998); others include implementation of Hopf-like models of outer hair cell active nonlinear amplification or attenuation of the signal, using discrete chips (Martignoli et al. 2007) or approximations of such models in silicon (Hamilton et al. 2008; Katsiamis et al. 2009).

3.2 Spiking cochlea architectures

The latest silicon cochlea designs include either cascaded or parallel architectures for the BM, inner hair cells, and the spiral ganglion cells with AER outputs (Abdalla and Horiuchi 2005; Fragniere 2005; Sarpeshkar et al. 2005; Wen and Boahen 2006; Chan et al. 2007; Liu et al. 2010a).

Our latest cochlea implementation (AER-EAR2) is a binaural cochlea intended for spatial audition and auditory scene analysis (Liu et al. 2010a). It integrates many features of former designs in a more user-friendly form. It uses cascaded second-order-sections (SOSs) (Fig. 2c) which drive inner hair cells, which in turn drive multiple ganglion cells with different spike thresholds. The resonance of individual sections can be adjusted by a local digital-to-analog converter (QDAC). This chip (Fig. 2d) includes a variety of features including a matched binaural pair of cochleas, on-chip digitally controlled biases, on-chip microphone preamplifiers, and open-sourced host software algorithms (jAER 2007). A bus-powered USB board enables easy interfacing to standard PCs for control and processing (Fig. 2e). An example response of AER-EAR2 to speech is shown in Fig. 2f.

3.3 Applications of cochleas

Applications of analog silicon cochleas in auditory tasks have been explored by different groups. The first experiments were done on extracting pitch and spatial location (Lazzaro et al. 1993; van Schaik and Shamma 2004). Because of the availability of AER cochleas, the focus has shifted to computations that capitalize on the precise timing of the cochlear events. This precise timing is useful for estimating interaural time difference (ITD) cues which can be used to localize sound sources (Chan et al 2007; Liu et al. 2010a) and for speech recognition tasks (Usyal et al. 2006). Estimating interaural time differences (ITDs) is expensive to compute using conventional approaches because of the high sample rate required on the analog microphone output. Another potential application is the use of these cochlear circuits as

preprocessors for cochlear implant systems (Georgiou and Toumazou 2005; Sarpeshkar et al. 2005).

4. Processing events

AER sensor events can be digitally processed by algorithms running on conventional hardware (jAER 2007; Delbruck 2008). The main characteristics of these algorithms are that they are event-driven and use the precise timing of the events. The focus here is on visual processing because these algorithms are more evolved.

The only prior work on interfacing AER chips with computers was the pioneering work on processing spikes from a silicon cochlea for speech recognition (Lazzaro et al. 1993). The recent availability of AER vision and audition sensors with good performance and user-friendly computer interfaces has given new life to this field. For example, recent applications of the dynamic vision sensor (DVS) include object tracking for commercial products (Litzenberger et al. 2006; Bauer et al. 2007; Gritsch et al. 2008); robotic applications (Delbruck and Lichtsteiner 2007; Conradt et al. 2009), fall detection of the elderly in assisted living (Fu et al. 2008), and in studying the development of topological connections (Boerlin et al. 2009).

Methods for digitally processing events have evolved into the following classes: *filters* that clean up the input to reduce noise or redundancy; *labelers* that assign additional meaning to the events such as contour orientation or direction of motion; and *trackers* that use events to track moving objects. The filters and labelers generally use one or several retinotopic maps of event times. These maps can be thought of as pictures of the most recent event times. The representation of events as software objects allows attachment of arbitrary annotation. As events are processed, some events are discarded, and remaining events are labeled with additional information. Instead of expanding the representation by expanding the number of cell types (as in cortical processing), the digital events are assigned increasing interpretation.

4.1 Low level visual feature extraction

As one example of this style of processing we will consider a low level feature extraction labeler that annotates the events in the input stream with additional interpretation. The software *DirectionSelectiveFilter* (Fig. 3) shows how orientation information is first extracted and then used to compute local optical flow vectors. These vectors are then integrated over space to compute global translational, radial, and tangential optical flows. The input shown in Fig. 3a is generated from a dynamic vision sensor (DVS) camera flying around a room (in this case handheld simulated flying) painted with black and white squares on the walls. The orientation labeler labels the ON and OFF DVS events with an additional 'orientation type' that signals their spatial angle of maximum correlation with past events in the spatial vicinity.
This labeler uses a topographic memory of past event times. Events generated by pixels along a moving edge of a particular orientation will be more correlated in time than events at other orientations. The orientation labeler tags each event using criteria about the maximum allowed correlation time within the receptive field. Events that pass the correlation test are output as orientation events. The correlation time is the average time difference between this event and the past ones stored in the region of the receptive field, with smaller time differences indicating stronger correlations.

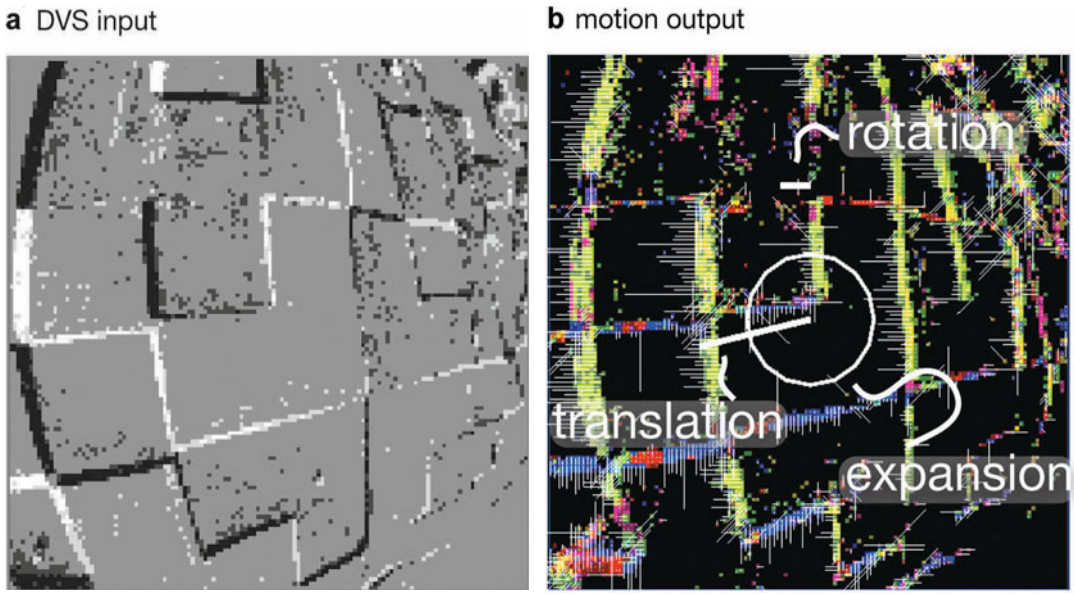


Fig. 3 Example output from the *DirectionSelectiveFilter* event labeler. **a** DVS input to labeler. Grey pixels indicate no events while white/black pixels indicate ON and OFF events caused by moving edges in the scene; **b** output of labeler. Local motion vectors (thin lines) are integrated over space and time to produce the global optical flow components for translation, rotation, and expansion (thick lines). Rotation and expansion vectors are centered on the center of the scene

The orientation events are then used to compute local motion vectors using the time of flight of orientation events. Each orientation event starts a search over the map of past orientation events in a direction normal to the orientation, and the most likely direction of motion is chosen based on a correlation measure similar to the one used for orientation. The speed is computed by the time of flight from the past orientation event. The output of the motion labeler (Fig. 3 b) shows how the local normal flow vectors (thin lines) result in estimates of the global translational flow (thick line from center), which in this case, points to the left. Similar global metrics of rotational and radial flow are also computed.

Low level methods for processing spikes like the *DirectionSelectiveFilter* could provide a means of bringing a hierarchical style of digital processing to AER events. But since the main success so far in practical applications is in tracking objects, we will now turn to this subject.

4.2 Tracking

Object tracking is widely-used visual capability. The main advantages of using an AER sensor for object tracking include system computational efficiency, low system latency, and the capability of the sensor to operate over a wide dynamic range of background intensities. One example is a tracker called the *RectangularClusterTracker* that tracks multiple moving objects (Litzenberger et al. 2006; Delbruck and Lichtsteiner 2007). This tracker treats an object as a spatially-connected source which generates a *cluster* of temporally correlated events. As objects move, they generate events which drag along the clusters. Clusters are spawned by events and disappear when moving clusters merge

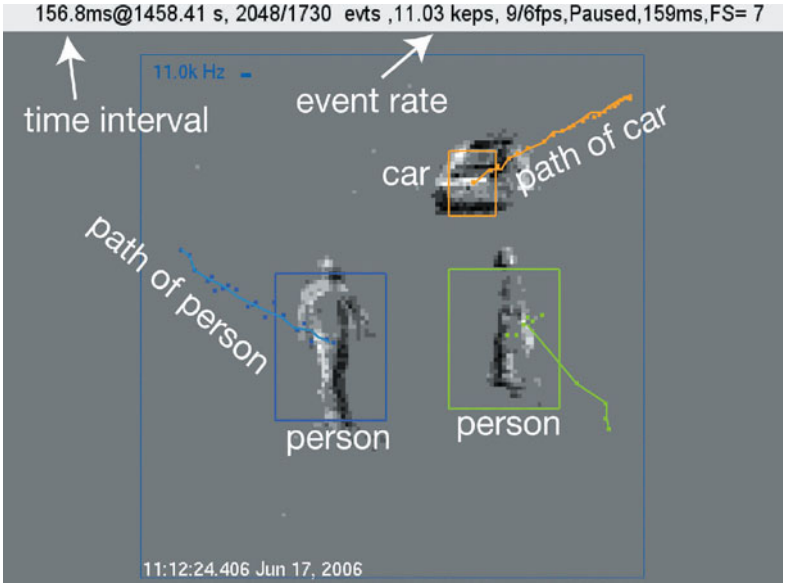


Fig. 4 Object tracking: RectangularClusterTracker tracks two persons and a car. The history of each cluster is shown as the dots

or when they are starved of events. Cluster sizes are determined either by knowledge of the scene (perspective) and object classes (e. g. cars), or they can be dynamic.

Tracking using events is advantageous because clusters do not have to be tracked over frames. Each pixel event updates the information about the size and location of its member cluster and only sensor pixels that generate events need to be processed. The primary processing cost is thus dominated by the search for the nearest existing cluster and the memory cost is limited only to the storage of the list of clusters.

Fig. 4 shows an example of how the *RectangularClusterTracker* can track a car and two people in a scene. Over the 157 ms window shown, the DVS produces 2048 events. 1730 events were left after pre-filtering to remove uncorrelated activity. The past locations of the clusters over 2048-event slices are shown by the dots. The sizes of the clusters dynamically come to equilibrium by balancing the events inside and surrounding the clusters.

The *RectangularClusterTracker* has been used in a robotic goalie that has a reaction latency of 3 ms with a 4 % processor load, using standard USB interfaces (Delbruck and Lichtsteiner 2007). This combination of metrics would be impossible to achieve using conventional frame based vision.

Other types of trackers have also been developed. For instance, the world's first pencil balancing robot uses a pair of dynamic vision sensors (DVSs) with embedded microcontrollers (Fig. 1f) to track the pencil in 3d space. The low memory and processing costs of events allow the pencil balancing robot to process input events at an effective frame rate of more than 10 kHz using only 200 mW embedded fixed-point microcontrollers (Conradt et al. 2009).

Work is ongoing to extend these notions about object tracking to include simultaneous tracking of multiple objects, object classification, and object recognition.

Summary and discussion

AER sensors are starting to make their way into usable high-performance prototypes. The nine papers from 2002 to 2010 on event-based sensors in the very competitive IEEE International Solid State Circuits Conference illustrate that this approach is starting to impact mainstream electronics (Barbaro et al. 2002; Fragniere 2005;

Mallik et al. 2005; Sarpeshkar et al. 2005; Wen and Boahen 2006; Lichtsteiner et al. 2006; Posch et al. 2007; Massari et al. 2008; Posch et al. 2010).

One achievement of this work is the development of event-based digital processing methods that capture the flavor of biological spike-based processing. These methods were only developed after sensors were built in a form that allowed their everyday use away from the lab bench. Although these methods have been developed as software algorithms on standard computers (jAER 2007), one can consider a variety of event-processing platforms ranging from conventional computers to the other extreme of using AER neuromorphic chips (Choi et al. 2005; Chicca et al. 2006; Serrano-Gotarredona et al. 2009). Our industrial partners are using an embedded DSP platform (Belbachir et al. 2007). We ourselves have used tiny microcontrollers (Conradt et al. 2009) and are starting to use field-programmable gate arrays (Linares-Barranco et al. 2007). This field has the potential for realization of small, fast, low power, embedded sensory-motor processing systems that are beyond the reach of traditional approaches under constraints of power, memory, and processor cost. The availability of event-based silicon retinas and cochleas will enable investigations of binding across visual and auditory sensory modalities using the fine temporal structure afforded by spikes. The asynchronous nature of AER data could inspire digital signal processing that breaks away from conventional regular-sample processing and that builds on top of decades of work in understanding computation by the nervous system (Liu and Delbruck 2010).

Acknowledgments

The authors gratefully acknowledge the constructive comments by the anonymous reviewer, F. Barth, M. Srinivasan, J. Harris, M. Rostogi, and N. da Costa.

References

Abdalla H, Horiuchi T (2005) An ultrasonic filterbank with spiking neurons, IEEE Intl. Symp. on Circuits and Systems (ISCAS 2005), pp. 4201–4204

Barbaro M, Burgi PY, Mortara A, Nussbaum P, Heitger F (2002) A 100 × 100 pixel silicon retina for gradient extraction with steering filter capabilities and temporal output coding. IEEE J. of Solid-State Circuits 37: 160–172

Bauer D, Belbachir AN, Donath N, Gritsch G, Kohn B, et al. (2007) Embedded vehicle speed estimation system using an asynchronous temporal contrast vision sensor. EURASIP J Embedded Syst 2007 (art. ID 82 174): 1–12

Belbachir AN, Litzenberger M, Posch C, Schon P (2007) Real-time vision using a smart sensor system. IEEE Intl Symp Industrial Electronics 2007. ISIE 2007, pp. 1968–1973.

Berner R, Lichtsteiner P, Delbruck T (2008) Self-timed vertacolor dichromatic vision sensor for low power face detection. IEEE Intl Symp Circuits and Systems (ISCAS 2008), pp. 1032–1035

Boerlin M, Delbruck T, Eng K (2009) Getting to know your neighbors: Unsupervised learning of topography from real-world, event-based input. Neural Computation 21: 216–238

Chan V, Liu S-C, van Schaik A (2007) AER EAR: A matched silicon cochlea pair with address event representation interface. IEEE Trans Circuits and Systems I: Regular Papers 54: 48–59

Chicca E, Whatley AM, Lichtsteiner P, Dante V, Delbruck T, et al. (2006) A multi-chip pulse-based neuromorphic infrastructure and its application to a model of orientation selectivity. IEEE Trans Circuits and Systems I: Regular Papers 54: 981–993

Choi TYW, Merolla PA, Arthur JV, Boahen KA, Shi BE (2005) Neuromorphic implementation of orientation hypercolumns. IEEE Trans Circuits and Systems I: Regular Papers 52: 1049–1060

Conradt J, Berner C, M., Delbruck T (2009) An embedded AER dynamic vision sensor for low-latency pole balancing. 5th IEEE Workshop on Embedded Computer Vision (in conjunction with ICCV 2009), Kyoto, Japan

Delbruck T, Lichtsteiner P (2007) Fast sensory motor control based on event-based hybrid neuromorphic-procedural system. IEEE Intl Symp Circuits and Systems (ISCAS 2007), pp. 845–848

Delbruck T (2008) Frame-free dynamic digital vision. Proc Intl Symp Secure-Life Electronics, Advanced Electronics for Quality Life and Society, pp. 21–26. Tokyo: University of Tokyo

Douglas R, Mahowald M, Mead C (1995) Neuromorphic Analog VLSI. Ann Rev Neurosci 18: 255–281

Fasnacht D, Delbruck T (2007) Dichromatic spectral measurement circuit in vanilla CMOS. IEEE Intl

Symp Circuits and Systems (ISCAS 2007), pp. 3091–3094

Fossum ER (1997) CMOS image sensors: electronic camera-on-a-chip. IEEE Trans Electron Devices 44: 1689–98

Fragniere E (2005) A 100-channel analog CMOS auditory filter bank for speech recognition. IEEE ISSCC Dig of Tech Papers, pp. 140–589

Fu Z, Delbruck T, Lichtsteiner P, Culurciello E (2008) An address-event fall detector for assisted living applications. IEEE Trans Biomed Circuits and Systems 2: 88–96.

Georgiou J, Toumazou C (2005) A 126-uW cochlear chip for a totally implantable system. IEEE J. Solid-State Circuits 40: 430–443

Gold B, Morgan N (2000) Speech and audio signal processing: John Wiley and Sons, Inc. New York, NY

Gritsch G, Litzenberger M, Donath N, Kohn B (2008) Real-time vehicle classification using a smart embedded device with a 'silicon retina' optical sensor. ITSC08, pp. 534–538. Bejing, China

Hamilton T, Tapson J, Jin CT, van Schaik A (2008) An active 2-D silicon cochlea. IEEE Trans Biomed Circuits and Systems 2: 30–43

Indiveri G, Liu S-C, Delbruck T, Douglas R (2009) Neuromorphic systems. In: LSquire (ed) Encyclopedia of neuroscience, pp. 521–528: Academic Press

jAER (2007) jAER Real time sensory-motor processing for spike based address-event representation (AER) sensors and systems available: http://jaer.wiki.sourceforge.net

Katsiamis A, Drakakis E, Lyon R (2009) A biomimetic, 4.5uW, 120+ dB, log-domain cochlea channel with AGC. IEEE J Solid-State Circuits 44: 1006–1022

Kramer J (2002) An ON/OFF transient imager with event-driven, asynchronous read-out. IEEE Intl Symp Circuits and Systems (ISCAS 2002), pp. 165–168

Lazzaro J, Wawrzynek J, Mahowald M, Sivilotti M, Gillespie D (1993) Silicon auditory processors as computer peripherals. IEEE Trans Neural Networks 4: 523–528

Lennie P (2003) The cost of cortical computation. Current Biology 13: 493–497

Lenoro-Bardallo JA, Serrano-Gotarredona T, Linares-Barranco B (2010) A spatial calibrated contrast AER vision sensor with adjustable contrast threshold. IEEE Intl Symp Circuits and Systems (ISCAS 2010), pp. 2426–2429

Lichtsteiner P, Posch C, Delbruck T (2006) A 128×128 120 dB 30 mW asynchronous vision sensor that responds to relative intensity change. ISSCC Dig Tech. Papers, pp. 508–509 (27.9). San Francisco

Lichtsteiner P, Posch C, Delbruck T (2008) A 128×128 120 dB 15us latency asynchronous temporal contrast vision sensor. IEEE J Solid State Circuits 43: 566–576

Linares-Barranco A, Gómez-Rodríguez F, Jiménez A, Delbruck T, Lichtsteiner P (2007) Using FPGA for visuo-motor control with a silicon retina and a humanoid robot. IEEE Intl Symp Circuits and Systems (ISCAS 2007), pp. 1192–1195

Litzenberger M, Posch C, Bauer D, Schön P, Kohn B, et al. (2006) Embedded vision system for real-time object tracking using an asynchronous transient vision sensor. IEEE Digital Signal Proc Workshop 2006, pp. 173–178. Grand Teton, Wyoming

Liu SC, Kramer J, Indiveri G, Delbruck T, Douglas R (2002) Analog VLSI: circuits and principles: MIT Press, Cambridge, MA

Liu SC and Delbruck T (2010) Neuromorphic sensory systems. Curr. Opin. in Neurobiol 20: 288–295

Liu SC, van Schaik A, Minch BA, Delbruck T (2010a) Event-based 64-channel binaural silicon cochlea with Q enhancement mechanisms. IEEE Intl Symp Circuits and Systems 2010 (ISCAS 2010), pp. 2027–2030

Liu SC, Mesgarani N, Harris, J, Hermansky, H (2010b) The use of spike-based representations for hardware audition systems. IEEE Intl Symp Circuits and Systems 2010 (ISCAS 2010), pp. 505–508

Lyon RF, Mead C (1988) An analog electronic cochlea. IEEE Trans Acoustics Speech and Signal Processing 36: 1119–1134

Mahowald MA (1992) VLSI analogs of neuronal visual processing: a synthesis of form and function. Computation and neural systems, Caltech, Pasadena, California

Mahowald MA (1994) An analog VLSI system for stereoscopic vision: Kluwer, Boston, MA

Mahowald MA, Mead C (1991) The silicon retina. Sci Am 264: 76–82

Mallik U, Clapp M, Choi E, Cauwenberghs G, Etienne-Cummings R (2005) Temporal change threshold detection imager. IEEE ISSCC Dig Tech. Papers, pp. 362–363

Martignoli S, van der Vyver J-J, Kern A, Uwate Y, Stoop R (2007) Analog electronic cochlea with

mammalian hearing characteristics. Applied Physics Letters 91 (064 108)

Massari N, Gottardi M, Jawed S (2008) A 100uW 64 × 128-pixel contrast-based asynchronous binary vision sensor for wireless sensor networks. IEEE ISSCC Dig Tech Papers, pp. 588–638

Mead C (1990) Neuromorphic electronic systems. Proc IEEE 78: 1629–1636

Olsson JAM, Hafliger P (2009) Live demonstration of an asynchronous integrate-and-fire pixel-event vision sensor. IEEE Intl Symp Circuits and Systems (ISCAS 2009), pp. 774–774

Pelgrom M, Tuinhout H, Vertregt M (1998) Transistor matching in analog CMOS applications. IEDM Tech Dig: 915–918

Posch C, Hofstatter M, Matolin D, Vanstraelen G, Schon P, et al. (2007) A dual-line optical transient sensor with on-chip precision time-stamp generation. IEEE ISSCC Dig Tech Papers, pp. 500–618

Posch C, Matolin D, Wohlgenannt R (2010) A QVGA 143 dB DR asynchronous address-event PWM dynamic image sensor with lossless pixel-level video compression. IEEE ISSCC Dig Tech. Papers, pp. 400–401

Posch C, Matolin D, Wohlgenannt R, Maier T, Litzenberger M (2009) A microbolometer asynchronous dynamic vision sensor for LWIR. IEEE Sensors Journal 9: 654–664

Rodieck R (1998) The first steps in seeing: Sinauer Associates, Sunderland, MA

Ruedi PF, Heim P, Gyger S, Kaess F, Arm C, et al. (2009) An SoC combining a 132 dB QVGA pixel array and a 32 b DSP/MCU processor for vision applications. IEEE ISSCC Dig Tech Papers, pp. 46–47

Ruedi PF, Heim P, Kaess F, Grenet E, Heitger F, et al. (2003) A 128 × 128, pixel 120-dB dynamic-range vision-sensor chip for image contrast and orientation extraction. IEEE J. Solid-State Circuits 38: 2325–2333

Sarpeshkar R (1998) Analog versus digital: Extrapolating from electronics to neurobiology. Neural Computation 10: 1601–38

Sarpeshkar R, Baker MS C., Sit JJ, Turicchia L, Zhak S (2005) An analog bionic ear processor with zero-crossing detection. IEEE ISSCC Dig Tech Papers, pp. 78–79

Sarpeshkar R, Lyon RF (1998) A low-power wide-dynamic-range analog VLSI cochlea. Analog Integrated Circuits and Signal Processing 16: 245–274

Serrano-Gotarredona R, Oster M, Lichtsteiner P, Linares-Barranco A, Paz-Vicente R, et al. (2009) CAVIAR: A 45 k neuron, 5 M synapse, 12 G connects/s AER hardware sensory–processing–learning–actuating system for high-speed visual object recognition and tracking. IEEE Trans Neural Networks 20: 1417–1438

Uysal I, Sathyendra H, Harris JG (2006) A biologically plausible system approach for noise robust vowel recognition. IEEE Proc Midwest Symp Circuits and Systems, pp. 245–249

van Schaik A, Shamma S (2004) A neuromorphic sound localizer. Analog Integrated Circuits and Signal Processing 39: 267–273

Watts L, Kerns DA, Lyon RF, Mead CA (1992) Improved implementation of the silicon cochlea. IEEE J. of Solid State Circuits 27: 692–700

Wen B, Boahen K (2006) A 360-channel speech preprocessor that emulates the cochlear amplifier. IEEE ISSCC Dig Tech Papers, pp. 556–557

Zaghloul KA, Boahen K (2004) Optic nerve signals in a neuromorphic chip II: Testing and results. IEEE Trans Biomed Engineering 51: 667–675

The mode-sensing hypothesis: matching sensors, actuators and flight dynamics

7

Holger G. Krapp, Graham K. Taylor, J. Sean Humbert

Contents

Abstract

Here we elaborate upon the recent hypothesis that the sensory systems of insects are matched to their flight dynamics, such that they are configured to make or encode measurements within a modal coordinate system. This hypothesis is inspired by several distinctive organizational principles of insect sensory systems: namely, that insects appear to be configured a) to sense relative, rather than absolute, quantities; b) to make measurements in highly non-orthogonal axis systems; and c) to fuse sensory inputs from different modalities prior to using them as feedback to the actuators. Having elaborated upon the hypothesis itself and considered the functional advantages of the resulting control architecture, we discuss some of the physiological details of how the requisite coordinate systems might in practice be set up in the fly visual system. We also provide a mathematical framework for testing the quantitative match between sensory system and flight dynamics in the specific context of the visual systems of flies.

Holger G. Krapp
Imperial College London, Department of Bioengineering, Royal School of Mines, South Kensington Campus,
London, SW7 2AZ, UK
e-mail: h.g.krapp@imperial.ac.uk

1. Introduction

The study of insect flight control and its applications has reached an important juncture. Engineers are making headway in exploiting algorithmic tools inspired by the mechanisms of insect flight control (Taylor and Krapp 2007), with progress greatly facilitated by the application of classical con-

trol theory to biological systems. However, we believe that the time has come to ask whether biological systems, and by extension bio-inspired systems, can really be configured in the same way as conventional control systems. Taylor and Krapp (2007) identified three organizational principles of insect sensory systems, which run counter to the principles by which engineered, as opposed to biological, systems are classically organized. Firstly, in common with the sense organs of most animals, the sensors involved in insect flight control are adapted to measure relative, rather than absolute, quantities, typically sensing disturbances rather than steady states. Secondly, insects are adapted to make these measurements in highly non-orthogonal, yet highly structured, axis systems. Thirdly, the descending neurons that convey sensory input downstream to the flight motor are adapted to encode composite quantities, combining inputs from different sensory modalities. Taylor and Krapp argued that these principles point to the existence of a fundamentally different type of control architecture in insects, which they explained by proposing the "mode-sensing hypothesis" (Taylor and Krapp 2007). This states that the sensory systems of insects are specifically matched to their modes of motion[1], and represents a particular, albeit hypothetical, use of the more general principle of matched filtering that is so widespread in the sensory systems of living organisms. Matched filters are expected to evolve so as to detect the most salient signals, so the mode-sensing hypothesis makes evolutionary sense, because the most salient motion components in flight are those associated with its modes of motion. The aim of this paper is to develop the mode-sensing hypothesis in the context of the visual system of flies. We begin by formalizing mode-sensing hypothesis (Section 2.). We then consider how mode-sensing might work in practice in the fly visual system (Section 3.). Finally, we provide a mathematical framework for testing the mode-sensing hypothesis (Section 4.).

2. Formalizing the mode-sensing hypothesis

We may write a linearized model of the flight dynamics of an insect as:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} \tag{1}$$

where **x** is the state vector and **A** is the closed-loop system matrix (Taylor and Thomas 2003). This model characterizes the rigid body dynamics without regard to the flapping dynamics of the wings, which is reasonable if the forcing frequency of the wingbeat is at least an order of magnitude higher than the highest frequency of any oscillatory rigid body modes. The elements of the state vector **x** denote small perturbations from equilibrium in the kinematic states of the system. For example, if we restrict ourselves to considering symmetric motion, then the state vector might include the component of velocity along the longitudinal body axis, the component of velocity along the dorso-

[1] The movement of a linear system can be described as a linear sum of its modes of motion. These may be either oscillatory (e. g. swinging of a pendulum) or monotonic (e. g. deceleration of a parachute), and either stable (i. e. tending to decay with time) or unstable (i. e. tending to grow with time). For example, the symmetric motions of an aircraft away from equilibrium typically involve a fast pitch oscillation, known as the short period mode, superposed with a slower change in speed, altitude, and pitch, known as the phugoid mode.

ventral body axis, the angular velocity about the pitch axis, and the pitch attitude of the longitudinal body axis with respect to the horizontal. The elements of the system matrix **A** determine how the system responds to perturbations.

The motion of a linear system is a linear combination of all of its dynamical modes, so it is reasonable to expect that the individual sensors of an insect will all have evolved to be strongly excited by at least one of these modes (Taylor and Krapp 2007). Moreover, if we know the characteristic pattern of self-motion that is associated with a particular mode, then we can predict precisely how the sensors should be tuned to maximize observability of that mode of motion. This information is contained in the right eigenvectors of the system matrix **A**, which characterize the relative phase and magnitude of changes in the state variables during excitation of the modes of motion. Because the axes of rotation and translation that are associated with the different modes are likely to be non-orthogonal, matching the sensors of a system to the relevant components of its right eigenvectors should enhance signal to noise ratio, providing a possible adaptive explanation of the non-orthogonal arrangement of insect sensors. Spatial tuning of the sensors to the right eigenvectors does not result in any obvious simplification to the underlying control system, but as we now show, considerable simplification might be achieved if the descending sensory input to the flight motor were also matched to the left eigenvectors of the system matrix (cf. Tarokh 1992) (see also Section 4.).

Eq. (1) is a coupled set of first order differential equations, which can be decoupled by transforming it into a particular basis, known as a modal coordinate system (Stevens and Lewis 2003). The modal coordinates are defined by the following transformation:

$$\xi = \mathbf{M}^{-1}\mathbf{x} \tag{2}$$

where **M** is a matrix of right eigenvectors of the system matrix **A**, called the modal matrix (Stevens and Lewis 2003). Notice, however, that it is the inverse of the modal matrix, $\mathbf{M}^{-1}$, that appears in Eq. (2), which can be shown to be a matrix of left eigenvectors of **A**. Mathematically, the left eigenvectors of a linear system are equivalent to the right eigenvectors of the transpose of the system matrix, and although they have no easy or intuitive physical interpretation, they are specifically related to the controllability of the system (Stevens and Lewis 2003). The modal coordinates (ξ) which they specify are weighted sums of the state variables contained in **x**, with the weightings for a given mode specified by its left eigenvector. Provided that the eigenvalues of **A** are distinct, Eq. (2) may be used to rewrite Eq. (1) as:

$$\dot{\xi} = \mathbf{M}^{-1}\mathbf{A}\mathbf{M}\xi \equiv \mathbf{J}\xi \tag{3}$$

where **J** is the diagonal matrix[2] of eigenvalues of **A**. This diagonalization means that Eq. (3) is a set of uncoupled first-order differential equations (Taylor and Thomas 2003), which represents a considerable simplification over the coupled set of equations in Eq. (1).

Because we tend to engineer sensing systems in an orthogonal way, the original state space representation of Eq. (1) is usually the most natural way to describe them. For a biological system, however, Eq. (3) might be

[2] If any of the eigenvalues are complex, then this matrix will contain complex entries on the diagonal. It can, however, be transformed into a real block diagonal matrix by using a different similarity transformation to define the modal coordinates. Specifically, columns of the modal matrix that contain a conjugate pair of eigenvectors should be replaced with columns containing the real and imaginary parts of one member of the conjugate pair.

the more natural representation: living organisms are rarely organized to be orthogonal in three-dimensional space, and a modal coordinate system that represents how the organism actually moves would seem to be a more natural frame of reference. We further hypothesise that input from the various individual sensors of insects is combined and weighted by the descending neurons so that the integrated signals sent to the flight motor are encoded in a modal coordinate system. The advantage of this would be a complete decoupling of the control of each mode, resulting in a parallel, rather than convergent, control architecture. This provides a possible adaptive explanation of why input from different sensory modalities is combined in the descending neurons. Furthermore, it should be possible in principle to use the left eigenvectors of the system to predict precisely how these inputs should be combined.

The mode-sensing hypothesis may therefore be summarized as two distinct but non-exclusive hypotheses. First, that the sensors of an insect are matched to the characteristic patterns of self-motion associated with its modes. These are specified by the right eigenvectors of the system. Second, that the input from the various sensors is fused so that the descending sensory input is encoded in a system of modal coordinates. These are specified by the left eigenvectors of the system. It is important to note that the left and right eigenvectors of a system are different, except in the special case that the system matrix is symmetric, so the modal coordinates are not related in any straightforward physical fashion to the combination of self-motions involved in a particular mode. In the next section, we focus upon the first hypothesis, asking whether there is any evidence that the visual system of flies is matched to their likely modes of motion.

3. Mode-sensing in practice

3.1 Estimating self-motion from optic flow: global patterns and local motion detection in insects

Movements of an animal relative to its environment result in global patterns of retinal image shifts, or optic flow fields (Gibson 1950). Optic flow is composed of local vectors indicating the direction and speed with which the environment moves at a given position across the animal's eye. Its global structure depends on the animal's rotational and translational self-motion components described in Cartesian coordinates (Koenderink and van Doorn 1987, Fig. 1a). While the magnitude of rotational optic flow only depends on angular rates, the relative magnitude of local, translational optic flow also depends on the distance between the eyes and objects in the environment. The latter is a crucial prerequisite for estimating the 3D-layout of the environment, relative ground speed, and distance flown (Srinivasan et al. 2001). Distance-invariant rotational optic flow is more important in the context of gaze stabilization and flight stabilization.

The global structures of rotational and translational optic flow are distinctly different. Velocity vectors in a translation flow field are aligned along great circles connecting two singularities where no relative motion occurs which are defined by the animal's linear flight trajectory (Fig. 1b). In a rotation-induced flow field all velocity vectors are aligned along parallel circles centered with singularities defined by the axis of rotation. Half way in between the singularities in a rotational flow field, the velocity vectors become maximal (Fig. 1c). Similarly, in translational optic flow fields maximal velocity vectors also appear between the singularities – but the distance-dependence of

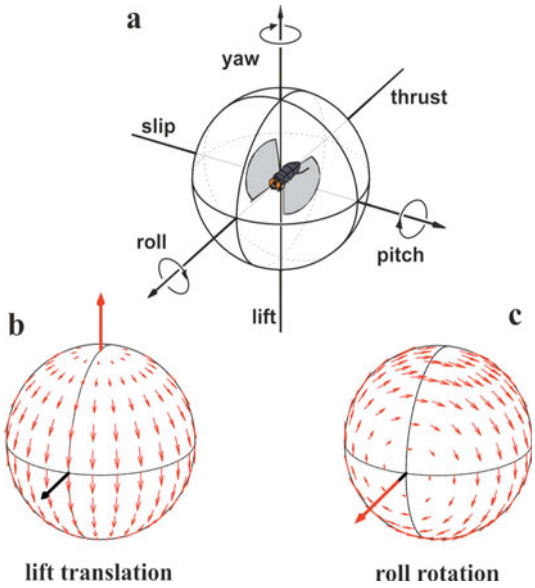


Fig. 1 Self-motion and optic flow. **a** The fly's self-motion components may be described in terms of three translations (thrust, sideslip, lift) and three rotations (roll, pitch, yaw) with respect to its cardinal body axes. **b** A translation results in an optic flow field where the local velocity vectors are aligned along great circles connecting the focus of expansion (red vertical axis) with the focus of extraction (opposite direction). Small red arrows indicate the direction and relative magnitude of local retinal image shifts. **c** Rotation induces an optic flow field where all velocity vectors are aligned along parallel circles centred with the axis of rotation (red horizontal axis; for roll rotation). Figure modified from Taylor and Krapp (2007)

translation-induced flow vectors may result in local deviations.

Analyzing optic flow patterns across its eyes should enable an animal to visually estimate its self-motion components (Koenderink and van Doorn 1987). There is a notorious problem, however, which all biological systems face: the ambiguity of local sensory information. The analysis of optic flow requires a mechanism that retrieves directional motion information. In most arthropods this is achieved by elementary movement detectors (EMDs) which perform a spatio-temporal correlation of light levels at neighbouring retinal positions (revs.: Reichardt 1961; Borst and Egelhaaf 1993; Egelhaaf and Borst 1993). There is experimental evidence in flies that directional motion processing mainly takes place along the orientation of ommatidial rows within the hexagonal compound eye lattice (Buchner 1984; Schuling et al. 1989; Petrowitz et al. 2000; Egelhaaf et al. 2002). Each EMD has a preferred direction, i. e. the direction of motion it is most sensitive to. Local directional information, however, is not sufficient to infer the self-motion component that has caused it. For instance, local downward motion in lateral visual field occurs during roll rotation but also during upward lift (cf. Fig 1 b and Fig 1 c). To overcome such ambiguity, the signals of those movement detectors whose preferred directions coincide with the direction of local velocity vectors in a specific flow field have to be spatially integrated across a sufficiently large area of the visual field.

3.2 Matched filters for optic flow: the lobula plate tangential cells

In blowflies a population of about 50–60 individually identified interneurons in the animal's third visual neuropile were found to respond to directional motion (revs.: Egelhaaf and Borst 1993; Hausen 1993; Borst and Haag 2002; Egelhaaf et al. 2002). These Lobula Plate Tangential Cells (LPTCs) receive input from thousands of EMDs which, altogether, cover most of the animal's spherical visual field. Each cell's receptive field organization is tuned to sense a specific optic flow field (Krapp and Hengstenberg 1996; Krapp et al. 1998, 2001; revs.: Krapp 2000; Taylor and Krapp 2007; Krapp and Wicklein 2008). Theoretical and experimental studies support the idea that many LPTCs function as matched filters for specific optic flow fields and provide the animal with robust

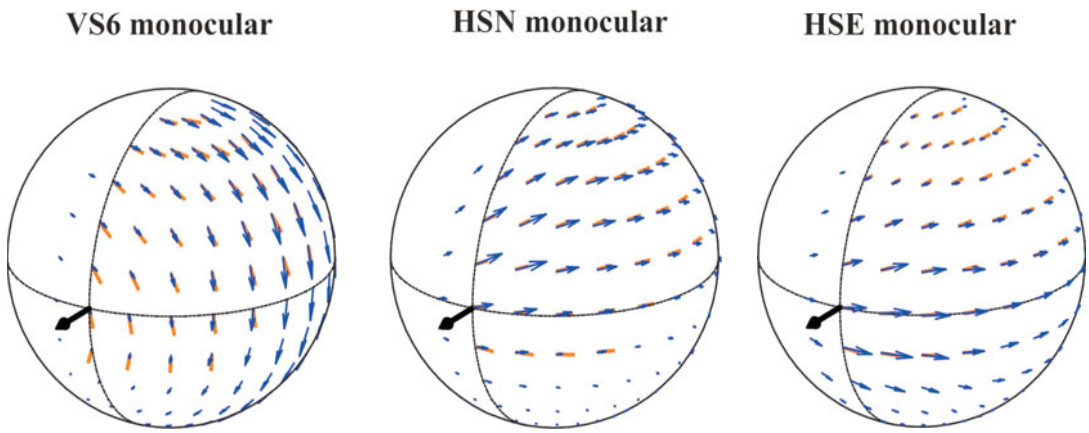


Fig. 2 Distribution of local motion preferences in three lobula plate tangential cells (LPTCs). **Left:** The orientation and length of each blue arrow indicates the local preferred direction and relative motion sensitivity of the VS6 cell. Note that the receptive field organization of VS6 resembles an optic flow field generated during roll rotation. Orange bars underneath the blue arrows show the local orientation of the v-rows in the fly's hexagonal eye lattice. **Middle:** Receptive field organization of the HSN cell. **Right:** Receptive field organization of the HSE cell. HS-cells respond to rotations around the yaw axis. During translational movements they are involved in estimating the 3D-layout of the surroundings. The results plotted were obtained in animals where the eye contralateral to the visual stimulus was occluded. VS = Vertical System; HSN = Horizontal System North(ern cell); HSE = Horizontal System Equatorial (cell). Data taken from Krapp and Hengstenberg (1996) and Krapp et al. (2001)

information about its self-motion (Franz and Krapp 2000; Dahmen et al. 2001; Karmeier et al. 2003, 2005; Franz et al. 2004) (see also chapter II.2 by S. Humbert et al.).

Here we will discuss the functional significance of two subgroups of LPTCs, the three Horizontal System cells (HS), (Hausen 1982), and the ten Vertical System cells (VS), (Hengstenberg 1982). The axon terminals of both HS- and VS-cells are connected to neck motor neurons in the head and to descending neurons conveying signals to the various motor systems in the animal's thorax. HS- and VS-cells are involved in flight and gaze stabilization (revs.: Hausen 1993; Hengstenberg 1993) and, as more recent work on HS-cells revealed, in visual distance estimation (Lindemann et al. 2005). HS-cells (Krapp et al. 2001) are well suited to signal rotations around the yaw axis and to indicate relative distance during phases of sideslip immediately after saccadic yaw turns (Lindemann et al. 2005; Karmeier et al. 2006). The cells' specificity to yaw rotations is increased by sophisticated ipsilateral and heterolateral network interactions between LPTCs which are mainly selective to horizontal wide-field motion (Haag and Borst 2001). VS-cells are tuned to sense rotations around horizontal body axes (Krapp et al. 1998). They are also interconnected, either through electrical synapses between nearest neighbours or spiking interneurons, which increases their tuning to specific rotations (Haag and Borst 2004). Figure 2 shows the receptive field organization of one VS- and two HS-cells plotted on top of the local orientation of ommatidial rows in the fly eye lattice.

The hypothesis that LPTCs function as matched filters for optic flow requires the template of local motion preferences, and thus their preferred self-motion specificity, to be invariant with respect to a wide range of operating conditions. Several studies over the last couple of years have shown that neither the adaptational state of the cells,

nor parameters such as pattern velocity and contrast, or the animal's locomotor state (Longden and Krapp 2009) alter the local preferred directions of the LPTCs (rev.: Taylor and Krapp 2007). It was also shown that the preferred rotation axis of an LPTC is not changed upon superposition of translational and rotational optic flow fields (Karmeier et al. 2003). Finally, the receptive field organization of LPTCs appears to be innate, rather than being shaped by early visual experience (Karmeier et al. 2001).

Although the distribution of the local motion preferences within the receptive field of LPTCs seems to be set in stone, their actual output signals are certainly not. Several non-linear dendritic- and contrast gain control mechanisms are in place to avoid saturation of the cell's signaling range (Borst et al. 1995; Harris et al. 2000; rev.: Taylor and Krapp 2007). Together with the fact that the inputting EMD signals are not linear in the velocity domain, such mechanisms prevent the cells' output from signaling true angular velocity. What they rather encode – in the context of estimating rotations – is information about the changes in angular rates around specific rotation axes (Taylor and Krapp 2007).

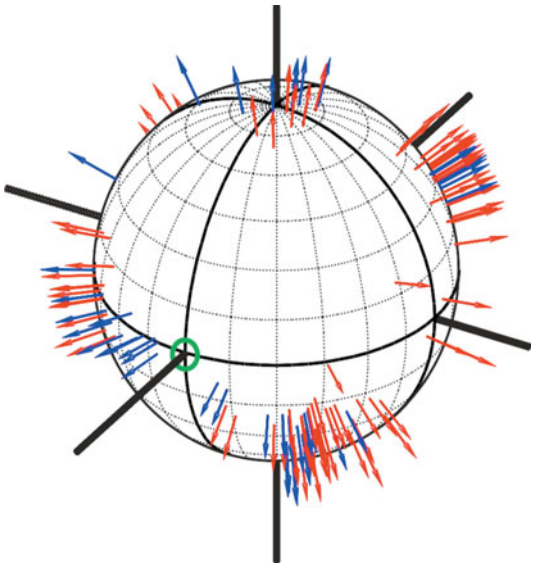


Fig. 3 Preferred rotation axes of LPTCs (blue) and neck motor neurons (red). Note that the axes' distributions overlap but are not identical. Deviations between the two populations of axes are partly due to the fact that neck motor neurons have a higher degree of binocularity, i. e. they receive stronger input from the contralateral eye. They may also reflect the transition from right to left eigenvectors describing measurement and control of modes of natural motion, respectively. Modified from Huston and Krapp (2008)

3.3 Sensorimotor coordinate transformation: what axes are covered by LPTCs?

LPTCs integrate signals obtained by elementary movement detectors within local retinal coordinates (Egelhaaf et al. 2002). Only recently, however, have the functional implications of this integration been explicitly spelled out (Taylor and Krapp 2007). As a result of the selective integration of local directional motion signals matched to a specific optic flow field, the output of the LPTCs is well suited to provide feedback to the various motor systems that control the animal's movements (Huston and Krapp 2008; Wertz et al. 2009). Thus, the LPTCs constitute the major building block of a sensorimotor coordinate transformation that converts signals obtained in sensory coordinates into a coordinate system that controls the gaze, and locomotor activities.

In the context of the mode-sensing hypothesis it is particularly interesting to identify which self-motion components are encoded by the set of LPTCs. Considering the bilateral symmetrical organization of the visual system, the 3 HS- and 10 VS-cells per lobula plate in *Calliphora* cover a total of 26 rotational degrees of freedom (Fig. 3, blue arrows). Thus, the dimensionality of the system by far exceeds a 3D orthogonal coordinate system. Why are self-motions so massively over-determined in the visuo-

motor pathway of the fly? In a noise-free measuring system, any self-motion component could be uniquely described by means of its projections into a set of three Cartesian axes. Any functional explanation must therefore take into account the limitations of robust biological sensing and the dynamical properties of the physical system to be controlled under closed-loop conditions.

A major constraint in biological sensing concerns noise induced, for instance, by the stochastic nature of opening and closing ion channels, which impairs neuronal information processing (White et al., 2000). The responses of VS-cells show a significant degree of variability when challenged with wide-field motion simulating rotations around various horizontal axes. The signal variance does not increase linearly with the response level – as it does in motion sensitive neurons in primates. Instead, it follows a bell-shaped function of response amplitude that peaks at values between maximum hyper- and depolarization relative to the cell's resting potential (Karmeier et al. 2005). Consequently, the signal-to-noise ratio of LPTC responses is best if they sense an optic flow field caused by a turn around their preferred rotation axis. By specifically aligning the LPTCs' preferred rotation axes with the right eigenvectors describing its modes of motion, the animal would increase the observability of these modes. The distribution of the preferred rotation axes in a cylindrical projection of VS-cells covering the right visual hemisphere (Krapp et al. 1998) reveals some interesting properties (Fig. 4). The data can be most readily interpreted by noting that a rotation about an axis at elevation 0° and azimuth 0° corresponds to a pure roll motion, a rotation about 0° elevation and 90° azimuth corresponds to a pure pitch motion, and that any non-zero elevation of the axis implies an additional component of yaw motion. Hence, VS1-VS2 may detect a banked turn involving yaw and pitch components, while VS3 may indicate the initial phase of re-assuming a level flight attitude, which necessarily involves a component of roll motion in addition to pitch and yaw (Taylor and Krapp 2007). Most interestingly, if the gaze is stabilized, the rotation axes of a subset of the VS-cells (VS4-VS10) are arranged in a way that suggests their sequential inhibition during a periodic Dutch roll mode, a lateral oscillatory motion involving coupled roll, yaw and sideslip (loc. cit.). On the other hand, if the gaze were to lead the azimuth of the body during a voluntary turn, then the same set of cells would tend to be excited sequentially (loc. cit.). At a first pass, therefore, the arrangement of the VS-cells appears to be consistent with the mode-sensing hypothesis (loc. cit.).

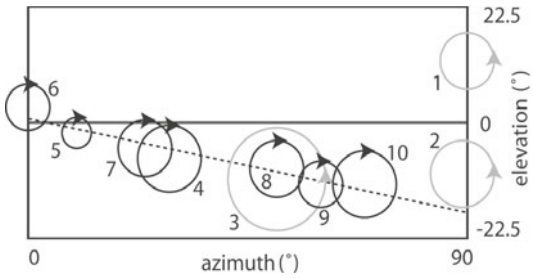


Fig. 4 Preferred rotational axes of the vertical system cells (VS1-VS10) of the visual system of *Calliphora* in a cylindrical projection of the right visual hemisphere. The circular arrows show the preferred axis and sense of rotation of each numbered cell, with diameter proportional to the standard deviation of the preferred axis across individuals. Cells in grey are excited by left-handed rotation about the axis shown; cells shown in black are excited by right-handed rotation. The opposite sense of rotation inhibits the cells. The dashed line is a regression of elevation on azimuth for the preferred axes of VS4-VS10 ($p < 0.05$, $R^2 = 0.93$). These cells are all stimulated by rotations of the same sense and will therefore be stimulated or inhibited sequentially in flight if the instantaneous axis of rotation moves along the line. Figure redrawn from Taylor and Krapp (2007)

Circumstantial evidence for the mode-sensing hypothesis comes from recent

studies on descending neurons (Wertz et al. 2009) and neck motor neurons (Huston and Krapp 2008) in *Calliphora*. Both these neuron types receive input from LPTCs and encode specific self-motion components. While descending neurons connect to locomotor systems in the thoracic ganglion, neck motor neurons control the fly gaze stabilization system that enables the animal to keep its head level during locomotion (loc. cit.). The preferred rotation axes of most visually driven motor neurons are arranged in a similar way to those of the LPTCs (Fig. 3, red arrows).

3.4 Integration of multisensory information

A problem in using visual information to estimate self-motion is that vision is notoriously slow. Blowflies encounter a dynamic stimulus range during flight that vastly exceeds the bandwidth within which the visual system could possibly encode changes in angular rates. The transduction process in the photoreceptors sacrifices speed for gain when photon energy is converted into receptor potentials (Hardie 1986). Furthermore, EMDs require extra time to retrieve directional motion information and are bound to cut off higher temporal frequencies to avoid aliasing which may result from a combination of high speed and high spatial frequency composition of the visual surroundings. Altogether, the visuomotor pathway operates on latencies which do not allow for closed-loop feedback control throughout the fly's entire flight envelope. It is probably for these reasons that flies employ a range of sensory modalities for state change estimation (Taylor and Krapp 2007). Besides the directional motion information provided by the compound eye, the fly possesses three simple lens eyes, ocelli. They are located on top of the head, measuring fast changes in flight attitude and are particularly sensitive to body rotations in the roll-pitch plane (revs.: Taylor and Krapp 2007; Goodman 1981). Taking advantage of a short pathway and high-speed signal conduction (Simmons et al. 1993) ocellar-induced compensatory head movements are significantly faster than those the compound eyes mediate (revs.: Hengstenberg 1991, 1993).

In addition to visual information several mechanosensory mechanisms provide proprioceptive, airflow and inertial information (Taylor and Krapp 2007) (see also Chapter IV.7 by T. Daniel et al.). They are faster than visual ones because of a direct transduction of mechanical energy into receptor potentials. However, like the halteres which measure Coriolis forces to retrieve angular rotation rates, most mechanosensory systems have a small gain in the low dynamic stimulus range (Nalbach 1994). Behavioural studies on fly gaze stabilization suggest that visual and mechanosensory mechanisms complement each other to cover the entire dynamic range of state changes flies encounter during locomotion (revs.: Hengstenberg 1991, 1993). The neuronal principles underlying the integration of signals obtained by different sensory mechanisms are currently under investigation. Depending on the specific pathway, modalities involved, and the level at which electrophysiological studies are carried out, both linear and non-linear combinations of sensory signals were found.

A recent study on blowfly ocelli has shown that their alternating illumination modulates the activity in some LPTCs (Parsons et al. 2006). In a follow-up study the activity of VS-cells were recorded intracellularly while stimulating the ocelli with spatiotemporal illumination patterns simulating attitude changes of the animal around horizontal rotation axes (Parsons et al. 2010). As for optic flow stimuli presented to the compound eyes, each VS-cell responded best when the ocellar stimulus mimicked a rotation around a

specific axis. While in both cases, the ocellar- and compound eye-mediated rotation tuning follows a cosine-shaped function there were two major differences: the delay of the ocellar responses was much shorter than that of the compound eye, but the spatial resolution at which VS-cells sample different horizontal rotations based on compound eye input was much higher than that for ocellar input. Upon ocellar input the cells encoded only three different rotation axes, i. e. 0°, +45°, and –45° azimuth, compared to nine different azimuthal axes for compound eye input. The different axis distributions for the two visual mechanisms were interpreted as a compromise between speed and accuracy. Fast but coarse information about attitude changes provided by the ocelli was thought to be complemented in an additive way by slow, but fine grained information on horizontal rotations from the compound eyes (loc. cit.). Electrophysiological studies on descending neurons in *Calliphora* receiving input from ocellar L-neurons support the idea of a linear superposition of signals from the ocelli and compound eyes (Haag et al. 2007).

Non-linear interactions between signals from different sensory mechanisms, however, have also been reported. Some neck motor neurons involved in fly gaze stabilization only generate a sufficient level of activity if the compound eyes are stimulated with directional motion *and* the halteres are moved (Huston and Krapp 2009). This finding seems to suggest a non-linear gating mechanism as proposed earlier for the flight control system (Heide 1983).

Obviously motion vision is not the only mechanism contributing to gaze and flight stabilization, but what makes it so important to flies? The motion vision pathway provided by the compound eyes is certainly the bottleneck in terms of processing speed, as ocelli and mechanosensory mechanisms mediate significantly faster feedback to the motor systems. However, no other modality allows the animal to establish a coordinate system in which self-motion components can be sensed in such a flexible and efficient way. Given that thousands of EMDs provide the input to individual LPTCs, matched filters to detect any combination of self-motion components could in principle be set up. The axes actually chosen, however, seem not to be random but facilitate the detection and ultimately the control of certain modes of motion. Once in place, the very same coordinates may be used by other sensory mechanisms to calibrate the signals used for actuation. In the next section we outline a control engineering framework that is based on optic flow processing and connects the properties of the LPTCs with the mode sensing hypothesis.

4. Control-theoretic models for visual motion decomposition by LPTCs

As discussed in Section 3.1., optic flow is the apparent visual motion produced by translation and rotation with respect to an environment. Patterns of optic flow therefore encode rigid body motions, and are also the primary source of relative distance to surrounding objects through translation-induced parallax (Gibson 1950). Koenderink and van Doorn (1987) expressed the optic flow $\dot{\mathbf{Q}}$ at a point $\mathbf{r}$ on a spherical imaging surface as:

$$\dot{\mathbf{Q}} = -\boldsymbol{\omega} \times \mathbf{r} - \mu\left[\mathbf{v} - \langle \mathbf{v}, \mathbf{r} \rangle \mathbf{r}\right]. \qquad (4)$$

Where $\boldsymbol{\omega}$ is the angular velocity vector and $\mathbf{v}$ the translational velocity vector. The function μ is called the nearness, and is equal to $\frac{1}{d}$, where d is the distance from the retina to the nearest object in the environment along the radial direction.

LPTCs respond with graded membrane potentials or shifts in spiking frequency whose polarity depends on the direction of motion (rev.: Hausen 1993). As described in Section 3., this response can be considered a comparison between a cell's preferred wide-field pattern of motion **F** (e. g., Fig. 2) and that of the visual motion stimulus $\dot{\mathbf{Q}}$, i. e. the instantaneous estimate of the optic flow pattern. Mathematically, one can model this operation as an inner product $\langle \dot{\mathbf{Q}}, \mathbf{F} \rangle$, interpreted as an abstraction of the *angle* between the patterns $\dot{\mathbf{Q}}$ and **F**. If the imaging surface is taken to be a sphere, the inner product takes the form:

$$\langle \dot{\mathbf{Q}}, \mathbf{F} \rangle = \int_{S^2} \dot{\mathbf{Q}} \cdot \mathbf{F} \; \mathrm{d}\Omega, \tag{5}$$

Where $d\Omega$ is the solid angle on the sphere (Tarokh, 1992). The output of a given LPTC with sensitivity pattern $\mathbf{F}_i$,

$$y_i = \langle \dot{\mathbf{Q}}, \mathbf{F}_i \rangle, \tag{6}$$

can then be computed analytically in terms of its dependence on proximity and translational and rotational velocity once a baseline (or parameterized family) of obstacle distributions μ is assumed. In general these computations might result in complicated nonlinear dependencies, however one can draw insight by examining the small signal approximation about a nominal flight condition. This results in a linear output equation of the form:

$$\mathbf{y} = C\mathbf{x}, \text{ where } C_{ij} = \left. \frac{\partial y_i}{\partial x_j} \right|_{\mathbf{x}_0}, \tag{7}$$

Here **x**, as in Eq. (1), is a vector whose entries are the perturbation states (δu,δq etc.) of the system, and $\mathbf{x}_0$ are their equilibrium values at the nominal flight condition. The matrix *C* is referred to as the observation matrix, with row *i* describing the dependence of the states **x** in the LPTC output y_i.

Given the above state-dependent model for LPTC outputs, it is now possible to describe a path forward to determine if the visual system of flies is matched to their likely modes of motion as described in Section 1. The observability measure for a given mode *i* can be defined as

$$m_i = |\delta_i| \left[e_i^* C^T C e_i \right]^{1/2}, \tag{8}$$

Where e_i is the right eigenvector associated with mode *i*, and δ_i is a constant complex scalar representing the distance of the e_i eigenvalue from the eigenvalues of the remaining modes (Tarokh, 1992). Upon further examination of this expression, it is clear that m_i decreases to zero as the angles of the right eigenvector e_i and the corresponding rows of the output matrix *C* become orthogonal. Hence, to maximize the observability of the overall system one would choose LPTC receptive field organizations such that the row vectors of the *C* matrix correspond to the eigendirections e_i, or restated, the motion preferences of the fly visual system are matched to their likely modes of motion.

To properly perform this analysis, flight dynamics models which include both the longitudinal and lateral-directional modes about hover and forward flight conditions are required. While there has been some recent progress in determination of the structure of the system matrix for dipteran flight modes about hover (Faruque and Humbert 2010a, 2010b), the absence of equivalent results for forward flight conditions prevents a comprehensive analysis at this point in time.

Conclusion

The hypothesis that insects are tuned to sense excitation of their modes of motion by making measurements within a modal coordinate system is, we believe, a natural way to interpret several otherwise puzzling aspects of the organization of insect sensory systems. Our hypothesis makes sense from the perspectives of evolution and ontogeny, because it is rea-

sonable to expect that the incremental nature of both processes will cause any sensory system functioning as a filter matched to some dominant motion component that the insect naturally experiences (i. e. matched to its right eigenvectors). The hypothesis also makes sense from a functional perspective, because matching the sensors to the dominant components of motion is obviously expected to improve the signal to noise ratio of each sensor. Moreover, a system configured to encode measurements within a modal coordinate system (i. e. matched to its left eigenvectors) would also benefit from a parallel, rather than convergent control architecture. Given that the eigenvectors refer to changes in state, it is also possible that such a system could make do with measurements of changes in state, rather than absolute state, and could thereby operate with greater bandwidth. In principle, the mode-sensing hypothesis might also apply to the arrangement of the sensory systems associated with other modes of locomotion, such as walking or swimming. Future work will test this hypothesis using the framework we have outlined above.

Acknowledgments

The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013)/ ERC grant agreement no. 204513 to GKT, the Biotechnology and Biological Sciences Research Council (BBSRC) Joint Grants BB/C007336/1 and BB/C518573/1 to HGK and GKT, US Air Force Research Laboratory Grant FA8655-09-1-3022 to HGK, and US Air Force Office of Scientific Research Joint Grants FA95500910075 and FA8655-09-1-3083 to JSH and HGK.

References

Borst A, Egelhaaf M (1989) Principles of visual motion detection. Trends Neurosci 12: 297–306

Borst A, Egelhaaf M, Haag J (1995) Mechanisms of dendritic integration underlying gain control in fly motion-sensitive interneurons. J Comput Neurosci 2: 5–1837

Borst A, Haag J (2002) Neural networks in the cockpit of the fly. J Comp Physiol A 188: 419–437

Buchner E (1984) Behavioural analysis of spatial vision in insects. In: Ali MA (ed) Photoreception and vision in invertebrates. Plenum Press, New York, pp. 561–621

Dahmen H, Franz M, Krapp HG (2001) Extracting ego-motion from optic flow: limits of accuracy and neuronal filters. In: Zanker JM, Zeil J (eds) Processing visual motion in the real world – A survey of computational, neural and ecological constraints. Springer Verlag, Berlin, Heidelberg, New York, pp. 143–168

Egelhaaf M, Borst A (1993) Movement detection in arthropods. Rev Oculomot Res 5: 53–77

Egelhaaf M, Kern R, Krapp HG, Kretzberg J, Kurtz R, Warzecha A-K (2002) Neuronal encoding of behaviourally relevant visual-motion information in the fly. Trends Neurosci 25: 94–100

Faruque I, Humbert JS (2010). Dipteran insect flight dynamics: Part 1: Longitudinal motions about hover. J Theo Biol 264: 538–552

Faruque I, Humbert JS (2010). Dipteran insect flight dynamics: Part 2: Lateral-directional motions about hover. J Theo Biol 265: 306–313

Franz MO, Krapp HG (2000) Wide-field, motion-sensitive neurons and optimal matched filters for optic flow. Biol Cybern 83: 185–197

Franz MO, Chahl JS, Krapp HG (2004) Insect inspired estimation of self-motion. Neural Comput 16: 2245–2260

Gibson JJ (1950) The perception of the visual world. Houghton Mifflin, Boston

Goodman L (1981) Organisation and physiology of the insect dorsal ocellar system. In: Autrum H (ed) Handbook of sensory physiology, Volume VII/6C, Springer Verlag, Berlin, pp. 201–181

Haag J, Borst A (2001) Recurrent network interactions underlying flow-field selectivity of visual interneurons. J Neuosci 21: 5685–5692

Haag J, Borst A (2004) Neural mechanism underlying complex receptive field properties of motion-sensitive interneurons. Nat Neurosci 7: 628–634

Haag J, Wertz A, Borst A (2007) Integration of lobula plate output signals by DNOVS1, an identified premotor descending neuron. J Neurosci 27: 1992–2000

Hardie RC (1986) The photoreceptor array of the dipteran retina. Trends Neurosci 9: 419–423

Harris RA, O'Carroll DC, Laughlin SB (2000) Contrast gain reduction in fly motion adaptation. Neuron 28: 595–606

Hausen K (1982) Motion sensitive interneurons in

the optomotor system of the fly I. The horizontal cells – structure and signals. Biol Cybern 45: 143–156

Hausen K (1993) Decoding of retinal image flow in insects. Rev Oculomot Res 5: 203–235

Heide G (1983). Neural mechanism of flight control in Diptera. In: Nachtigall W (ed) BIONA report, Vol. 2, Akad Wiss Mainz, Gustav Fischer, Stuttgart, New York, pp. 35–52

Hengstenberg R (1982) Common visual response properties of giant vertical cells in the lobula plate of the blowfly *Calliphora*. J Comp Physiol A 149: 179–193

Hengstenberg R (1991) Gaze control in the blowfly *Calliphora*: A multisensory two-stage integration process. The Neurosciences 3: 19–29

Hengstenberg R (1993) Multisensory control in insect oculomotor systems. Rev Oculomot Res 5: 285–298

Huston SJ, Krapp HG (2008) Visuomotor transformation in the fly gaze stabilization system. PLoS Biol 6: 1468–1478

Huston SJ, Krapp HG (2009) Non-linear integration of visual and haltere inputs in fly neck motor neurons. J Neurosci 29: 13 097–13 105

Karmeier K, Tabor R, Egelhaaf M, Krapp HG (2001) Early visual experience and the receptive field organization of optic flow processing interneurons in the fly motion pathway. Vis Neurosci 18: 1–8

Karmeier K, Krapp HG, Egelhaaf M (2003) Robustness of the tuning of fly visual interneurons to rotatory optic flow. J Neurophysiol 90: 1626–1634

Karmeier K, Krapp HG, Egelhaaf M (2005) Population coding of self-motion: applying Bayesian analysis to a population of visual interneurons in the fly. J Neurophysiol 94: 2182–2194

Karmeier K, van Hateren JH, Kern R, Egelhaaf M (2006) Encoding of naturalistic optic flow by a population of blowfly motion-sensitive neurons. J Neurophysiol 96: 1602–1614

Koenderink J, van Doorn A (1987) Facts on optic flow. Biol Cybern 56: 247–254

Krapp HG, Hengstenberg R (1996) Estimation of self-motion by optic flow processing in single visual interneurons. Nature 384: 447–468

Krapp HG, Hengstenberg B, Henstenberg R (1998) Dendritic structure and receptive-field organization of optic flow processing interneurons in the fly. J Neurophysiol 79: 1902–1917

Krapp HG (2000) Neuronal matched filters for optic flow processing in flying insects. Int Rev Neurobiol 44: 93–120

Krapp HG, Hengstenberg R, Egelhaaf M (2001) Binocular input organization of optic flow processing interneurons in the fly visual system. J Neurophysiol 85: 724–734

Krapp HG, Wicklein M (2008) Central processing of visual information in insects. In: Basbaum AI, Kaneko A, Shepherd GM, Westheimer G. (eds) The senses: A comprehensive reference. Vol. 1, Vision I, Masland R, Albright TD (eds) Academic Press, San Diego, pp. 131–204

Lindemann JP, Kern R, van Hateren JH, Ritter H, Egelhaaf M (2005) On the computations analyzing natural optic flow: quantitative model analysis of the blowfly motion vision pathway. J Neurosci 25: 6435–6448

Longden KD, Krapp HG (2009) State-dependent receptive field properties of optic flow processing interneurons. J Neurophysiol 102: 3606–3618

Nalbach G (1994) Extremely non-orthogonal axes in a sense organ for rotation: behavioural analysis of the dipteran haltere system. Neurosci 61: 155–163

Parsons MM, Krapp HG, Laughlin SB (2006) A motion-sensitive neurone responds to signals from the two visual systems of the blowfly, the compound eyes and ocelli. J Exp Biol 209: 4464–4474

Parsons MM, Krapp HG, Laughlin SB (2010) Sensor fusion in identified visual interneurons. Curr Biol 20: 624–628

Petrowitz R, Dahmen H, Egelhaaf M, Krapp HG (2000) Arrangement of optical axes and the spatial resolution in the compound eye of the female blowfly. J Comp Physiol A 186: 737–746

Reichardt W (1961) Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In: Rosenblith WA (ed) Principles of sensory communications. Wiley, New York, pp. 303–317

Schuling FH, Mastebroek HAK, Bult R, Lenting BPM (1989) Properties of elementary movement detectors in the fly *Calliphora erythrocephala*. J Comp Physiol A 165: 179–192

Simmons PJ, Jian S, Rind FC (1993) Responses *in vivo* to light signals by large, 2nd-order ocellar neurons of the blowfly, *Calliphora erythrocephala*. J Physiol Lond 473: 244–244

Srinivasan MV, Zhang S, Chahl JS (2001) Landing strategies in honeybees, and possible applications to autonomous airborne vehicles. Biol Bull, 200: 216–221

Stevens BL, Lewis FL (2003) Aircraft control and simulation, 2nd edn. Hoboken, Wiley, New York

Tarokh M (1992) Measures of controllability, observability, and fixed modes. IEEE Trans Autom Control 37: 1268–1273

Taylor GK, Thomas ALR (2003) Dynamic flight stability in the desert locust *Schistocerca gregaria*. J Exp Biol 206: 2803–2829

Taylor GK, Krapp HG (2007) Sensory systems and flight stability: What do insects measure and why? Adv Insect Physiol 34: 231–316

Wertz A, Gaub B, Plett J, Haag J, Borst A (2009) Robust coding of ego-motion in descending neurons of the fly. J Neurosci 29: 14993–15000

White JA, Rubenstein JT, Kay AR (2000) Channel noise in neurons. Trends Neurosci 23: 131

Adaptive encoding of motion information in the fly visual system

8

Rafael Kurtz

Contents

Abstract

Adaptation is a ubiquitous mechanism by which sensory cells and neurons match their response properties to the currently prevailing features of their input stimuli. Visual motion-sensitive neurons in the fly brain have been used as a valuable model system for the *in vivo* analysis of physiological mechanisms underlying neuronal adaptation. In this model system the functional significance of adaptation for visual control of flight movements can be evaluated. Although the effects of adaptation on neuronal sensitivity can in principle be understood as modifications of input-output functions, the exact nature of these changes is often unclear. In this review it is examined in how far the effects of adaptation with different stimulus parameters can be explained by different schemes of adaptation. One important conclusion from studies of adaptation in the fly motion vision system with simple stimulus paradigms as well as with complex, behaviorally generated stimuli is that adaptation improves the sensitivity for novel stimuli during exposure to sustained stimulation. Neuronal adaptation might thus facilitate important tasks such as object detection and obstacle avoidance during flight.

Rafael Kurtz
Bielefeld University, Department of Neurobiology
Post Box 100131, 33501 Bielefeld, Germany
e-mail: rafael.kurtz@uni-bielefeld.de

1. Introduction

In biological systems visual motion plays an important role in the control of locomotion and navigation. Thus it is not surprising that principles of visual motion computation have been adopted from biology in the design of technical systems, e.g. autonomously navigating robots. One prominent example is flight control, in particular course stabilization and obstacle avoidance during flight. Control of flight movements presents an exceptionally challenging task for visual guidance. Compared to movement in a plane, greater degrees of freedom have to be handled and flight stabilization has to operate sufficiently quickly and reliably, in particular when ground crashes present a risk. However, the sensors involved in visual flight control are confronted with one major problem: the range of intensities and dynamics that have to be handled might often exceed a sensor's instantaneous operating range. Neuronal adaptation has been proposed to provide a valuable means to self-calibrate the system to the currently prevailing range of stimulus intensities. Moreover, adaptation might optimize the system's ability to extract from its complex continually varying input those stimulus features that are particularly relevant in the current behavioral situation. Adaptation-induced modifications of signal processing have been shown to be relevant in many sensory systems across a large range of model organisms (reviews: Menini 1999; Eatock 2000; Benda and Herz 2003; Fettiplace and Ricci 2003; Matthews and Reisert 2003; Kohn 2007; Wark et al. 2007). In this chapter visual motion processing in the fly brain is taken as an example in which the functional significance of adaptation can be studied *in vivo* by physiological experiments.

2. Fly motion vision as a model for rapid sensory-motor control and for the transfer of biological principles to technical systems

The transfer of biological design principles to technical systems has been inspired by the in-depth study of sensory-motor control in several biological model systems (review: Webb 2002). One of these biological models turned out to be of exceptional value: visual motion-processing in flies (in particular the fruit fly *Drosophila*, blowflies and hoverflies). The significance of the fly as a model system for rapid sensory-motor control is due to a number of advantages. In flies, it is feasible to combine the analysis of visually guided locomotor behavior with the investigation of the neuronal architecture and the neuronal computations underlying this behavior (reviews: Egelhaaf et al. 2005; Egelhaaf 2008; Borst 2009). Unlike in many other model systems, in flies there appear to be comparatively few processing steps involved in the extraction of specific motion information from visual cues. Moreover, a prominent class of neurons, consisting of less than 100 cells per brain hemisphere, converts visual motion information into neuronal signals suitable for visually guided motor control. This class of neurons, the tangential cells (TCs) of the lobula plate, has been particularly well studied in the blowfly *Calliphora vicina* (Fig. 1A; reviews: Borst and Haag 2002; Egelhaaf et al. 2005). Particular TCs have distinct morphological and functional properties, making them individually identifiable in electrophysiological recording experiments (reviews: Hausen and Egelhaaf 1989; Borst and Haag 2002) and in functional imaging studies (review: Kurtz et al. 2008).

TCs perform a crucial step in the transfer of visual motion input into output signals for locomotor control: they integrate local mo-

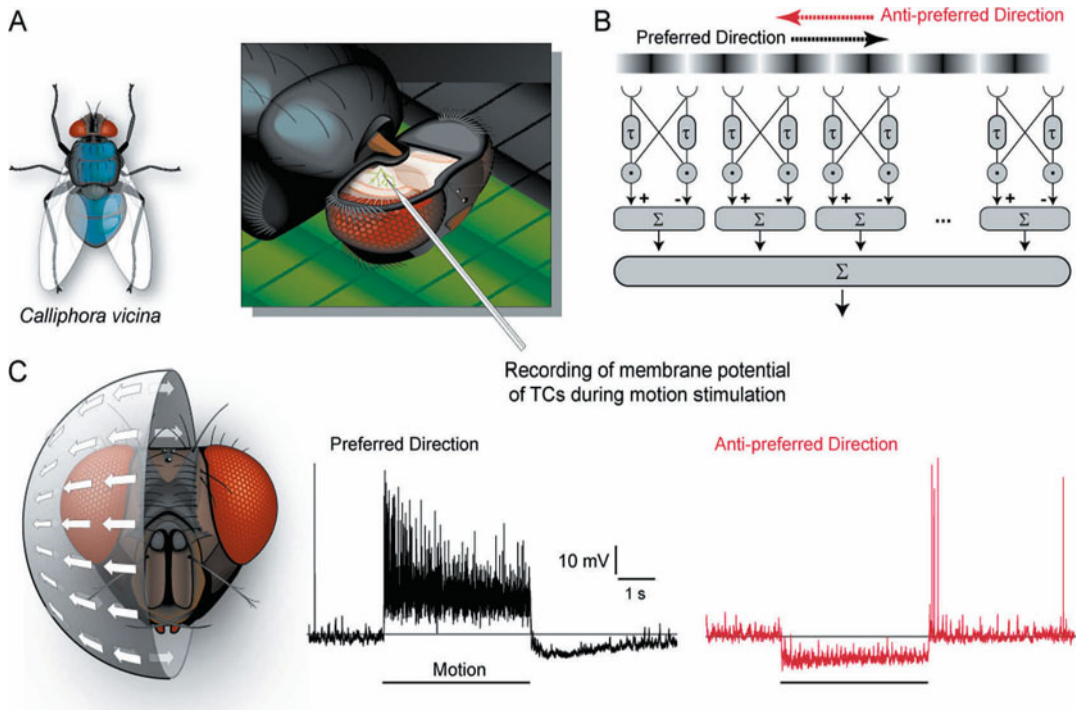


Fig. 1 Neuronal processing of visual motion in the fly brain. **A** Preparation of *Calliphora vicina* (left) used to record from TCs in the third visual neuropil, the lobula plate (right). **B** Computational principle thought to underlie motion computation. Motion is first computed by elementary motion detectors (EMDs), which correlate (by multiplication, "·") the brightness signal from one location in the visual field with a temporally delayed (by a low pass filter with time constant τ) brightness signal from a neighboring location. Subtraction of the output from one detector half unit from its mirror-symmetric counterpart yields a local motion signal that is positive for motion in preferred direction and negative for motion in anti-preferred direction. Integration over an array of EMDs provides global motion signals. TCs integrate at their dendrites the outputs from large 2-dimensional arrays of retinotopically arranged EMDs. **C** Motion response of a TC, consisting of a pronounced graded change of the axonal membrane potential and a change in the frequency of action-potential-like transients with variable amplitude ("spikelets"). Horizontal line indicates resting potential

tion inputs over large parts of the visual field (Krapp et al. 1998; Spalthoff et al. 2010). Reliable information about global motion patterns is extracted from retinal image shifts during self-motion (Krapp and Hengstenberg 1996). Thus, fly TCs are likely to play a prominent role in providing the motor system with the specific sensory signals that are required for visually guided flight stabilization and course control (see also Chapter II,5 by Krapp and Taylor).

3. Adaptation of sensory and neuronal systems to instantaneous stimulus levels

In biological as well as in technical systems, sensors face one common problem: The wide range of intensities over which their input signal may modulate is contrasted with a limited range of possible output states. Severe constraints to the working range of

sensory cells and neurons may result directly from the reversal potentials of the excitatory and inhibitory ionic currents or from a limit in spike rate. Corresponding working range limitations in technical systems are, for example, the finite capacity of photon conversions on a CCD chip or the mechanical displacement limit of an airflow sensor. The working range of biological as well as technical systems can be defined as the range over which a change in input intensity results into a sizable change of the output signal. In principle, the working range of a sensor could be broadened even in the presence of a fixed output limit by decreasing the system's gain, i. e. the slope of the input-output function. Unfortunately, this simple strategy is problematic because a lower slope of the input-output function inevitably results into weaker input-driven modulations of the output. Weak output modulations are, however, much more prone to corruption by downstream noise, in particular when fast modulations of the input signal need to be encoded, as is the case in visual motion processing.

Instead of reducing the slope of the input-output function, the entire function may be shifted horizontally to a certain range of input intensities, without changing the shape of the function. If such shifts become effective in a stimulus-history dependent way, the capability of a sensor to encode a large spectrum of possible input intensities can be improved without sacrificing the precision of output signals (Fig. 2A). In this way the system can take advantage of the fact that in most natural input signals characteristic temporal correlations are present, and use a strategy of "predictive coding" (see Srinivasan et al. 1982 for similar considerations concerning spatial visual coding). For example, the probability that luminance changes abruptly from bright sunlight to pitch dark is usually very low during natural visual stimulation. Thus the mean stimulus intensity in a certain time window of stimulus history often provides a good estimate of near future stimulus intensities. Therefore it might be useful to adjust a sensor's input-output relationship according to this estimate, thereby enhancing sensitivity to changes in the current stimulus intensity. When schematized as a single sigmoid input-output function a theoretically plausible adaptation-induced modification would shift the function such that the range of current stimulus intensities is optimally covered by the high-slope region of the curve (Fig. 2A).

In the fly motion vision system a large variety of different adaptation phenomena has been demonstrated over the past three decades (Srinivasan and Dvorak 1979; Maddess and Laughlin 1985; de Ruyter van Steveninck RR et al. 1986; Brenner et al. 2000; Harris et al. 2000; Kurtz et al. 2000; Fairhall et al. 2001; Reisenman et al. 2003; Borst et al. 2005; Heitwerth et al. 2005; Neri and Laughlin 2005; Kurtz 2007; Neri 2007; Kalb et al. 2008a; Kalb et al. 2008b; Liang et al. 2008; Kurtz et al. 2009a; Kurtz et al. 2009b; Nordström and O'Carroll 2009). In the following I will first outline the computational principle by which motion information is thought to be extracted from the visual stimulus. I will then describe some of the adaptation phenomena observed in visual motion-processing and examine the extent to which simple adaptation models (see Fig. 2) may help us to understand these phenomena. Finally, I will address the putative functional significance of motion adaptation under real-life conditions.

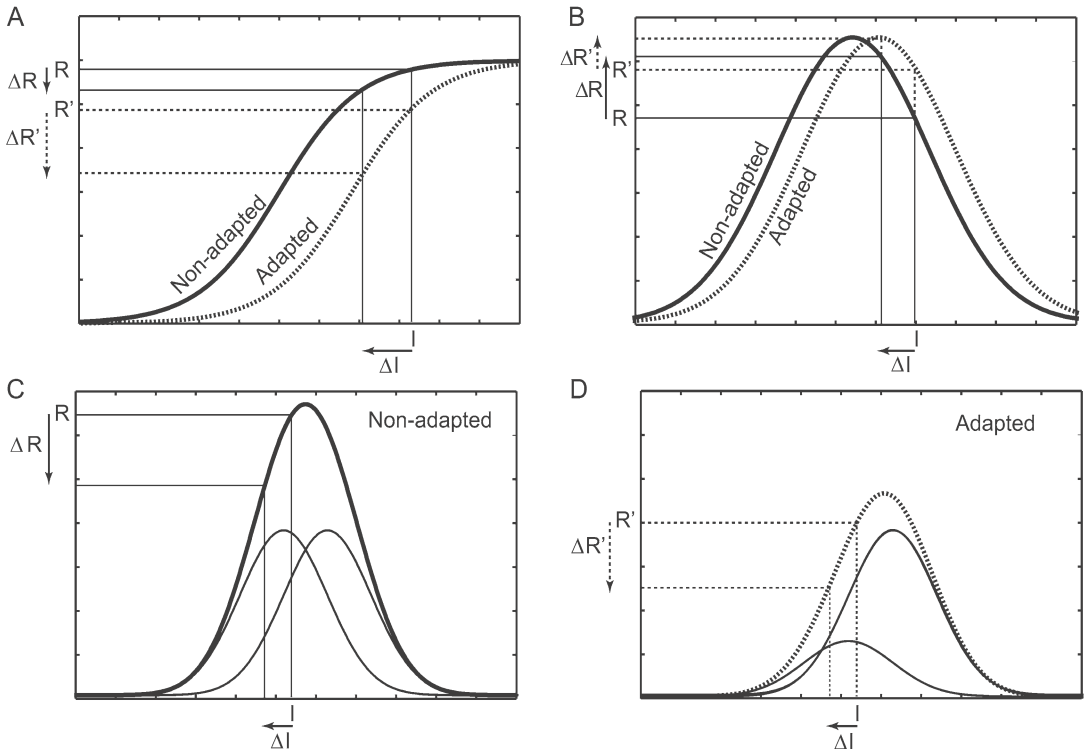


Fig. 2 Possible adaptation-induced changes of input-output functions. **A** Sigmoid input-output function, as is characteristic for the stimulus-response relationship of many neurons and sensory. Sustained high input intensity (I) causes adaptation, resulting in a lateral shift of the function towards higher input intensities. The shift causes attenuation of steady-state responses (R' vs R) and enhanced responses (ΔR' vs ΔR) to changes in input intensity (ΔI). **B** Input-output relationships may also follow bell-shaped functions, for example responses of visual motion-sensitive neurons to different velocities. Adaptation-induced shifts similar to those in *A* would enhance ΔR only in that part of the curve which has a positive slope. In contrast, as illustrated in the example, for inputs falling on parts of the curve with negative slope adaptation would lead to an increase of the steady-state response (R' vs R) and to a decrease of the response to changes in input intensity (ΔR' vs ΔR). **C**, **D** A bell-shaped input-output function (thick lines) is depicted, which results from integration of two inputs with peaks at different stimulus intensities (thin lines). Adaptation is assumed here to cause a simple attenuation of inputs (instead of a lateral shift as in **A** and **B**). This attenuation is "stimulus-specific", affecting only the input that shows the highest activity at the given stimulus level (I). This form of adaptation leads to a decrease in the steady-state value of the integrated response (R' vs R), but to nearly unaltered responses to changes in input intensity (ΔR' vs ΔR). Thus, the relative sensitivity to changes in input intensity (ΔR relative to R or, respectively, ΔR' relative to R') is enhanced with adaptation

4. The computational principle of visual motion detection and its implications for the dynamic characteristics of neuronal motion signals

During movement of an animal's or human's eyes, head or entire body, its retinas experience a continual image displacement. In a wide range of animal species, ranging from insects to monkeys, such global motion is processed by neurons that sample local motion across the visual field in a retinotopic way (reviews: Krapp 2000; Lappe 2000). In flies, this type of neuron is represented by the class of TCs (see above; reviews: Hausen and Egelhaaf 1989; Borst and Haag 2002; Egelhaaf et al. 2005). The responses of these neurons, however, depend not only on image velocity, but also on the contrast, the spatial frequency content and the orientation of pattern elements. Thus it is not surprising that adaptation of TCs has been shown to alter responsivity not only for velocity, but also for other features that characterize the motion stimulus (see sections 5–7).

As in other species, the local motion-detecting elements, which supply input to fly TCs, have so far largely evaded a direct experimental investigation. However, plausible computational models of local motion detection have been inferred based on the properties of TCs. One such model is the correlation-type motion detector, often referred to as the elementary motion detector (EMD) (review: Borst and Egelhaaf 1989). EMDs are based on correlating the appropriately filtered brightness signals from neighboring points in visual space (Fig. 1B). It has been shown that EMDs can explain the responses of motion-sensitive neurons to a wide range of motion stimuli (Egelhaaf and Reichardt 1987; Lindemann et al. 2005).

EMDs have several distinguishing features that are relevant in the context of motion adaptation. Most importantly, EMDs are not veridical sensors of local retinal velocities. (1) The time-averaged response to constant velocity motion increases with pattern velocity only within a certain velocity range; it then reaches an optimum and decreases again (Egelhaaf and Borst 1989). (2) The responses of individual EMDs modulate over time depending on the local pattern features within their receptive fields (Egelhaaf et al. 1989). (3) These modulations can be eliminated to some extent by spatial integration (Egelhaaf and Borst 1989). However, even spatially pooled EMD outputs have a characteristic phasic-tonic time course during stimulation with constant velocity (see Fig. 1C, 3). This temporal response profile is a direct consequence of the computational principle of EMDs and is thus manifested without an activity-dependent change in a system parameter, i. e. adaptation. Intriguingly, it has been shown that even more complex changes in response properties are emergent properties of the EMD. In particular, both in recordings from TCs as well as in EMD models with fixed parameters, it was observed that the slope of the neuronal input-output function changes when random velocity fluctuations of different modulation depths are presented (Borst et al. 2005). This phenomenon was termed "adaptation without parameter change" to differentiate it from "adaptation" in a strict sense, for which genuine changes in system parameters have to be present. In this context it is important to mention that for most of the adaptation phenomena described in the following it is at present not clear whether they result from physiological changes in the TCs themselves or from the properties of neurons at earlier stages of the visual pathway.

5. Motion adaptation changes the neuronal representation of image velocity and other stimulus parameters

In fly TCs, motion adaptation was first demonstrated in the H1-neuron (Maddess and Laughlin 1985). During adaptation with sustained pattern motion in preferred direction the response of H1 decreased. Nonetheless, the response transients elicited by velocity discontinuities, i. e. brief increments or decrements from baseline velocity, became more pronounced in the course of adaptation (see also Fig. 3 A). This result is in accordance with the view that adaptation improves neuronal sensitivity to changes in stimulus intensity around the current level. The mechanism underlying this phenomenon was proposed to be an activity-dependent down regulation of the time constant of the EMD low-pass filter. This parameter change would shift velocity tuning towards higher values (de Ruyter van Steveninck RR et al. 1986; Clifford and Langley 1996; Clifford et al. 1997). However, with such a mechanism alone, an enhancement of response transients to velocity discontinuities by motion adaptation would be restricted to low velocities. At high velocities the response transients evoked by changes in velocity might even become weaker in the adapted than in the non-adapted state. This prediction can be made with regard to the bell-shaped signature of velocity tuning curves (Fig. 2 B), but was not experimentally tested (Maddess and Laughlin 1985). Therefore, recently the question was addressed, whether enhanced sensitivity to velocity changes in the course of motion adaptation is also present at high baseline velocities (Kurtz et al. 2009 b). It was found that the transient deflections in spike rate of H1 in response to discontinuities in motion velocity were enhanced by

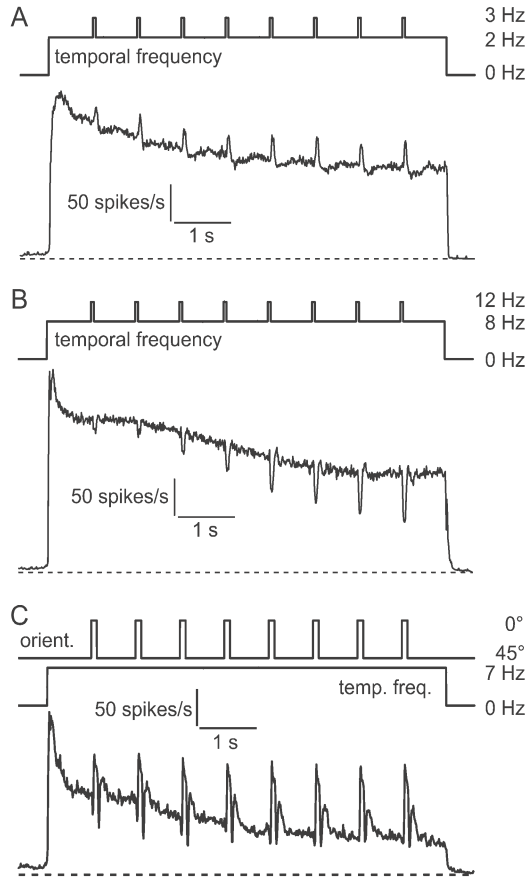


Fig. 3 Enhancement of responses to sudden stimulus changes with motion adaptation. **A** Spike rate of a *Calliphora* H1 neuron during motion of a periodic grating at constant velocity interspersed with brief increases in velocity. Although the neuronal response to the baseline velocity decreases with adaptation, the response to velocity increments is enhanced. The values in Hz indicate temporal frequency of the grating, which is linearly related to velocity. **B** Data obtained from a similar experiment but with baseline velocity above the steady-state optimum of the H1 neuron. Under this condition, increments in motion velocity lead to decrements in spike rate. Nevertheless, the amplitude of these decrements increases with adaptation. **C** Discontinuities in the motion stimulus produced by changing the orientation of the grating, and thus the direction of motion (0° = preferred direction). Again, adaptation enhances responses to the stimulus discontinuities. Modified from Kurtz et al. (2009 b)

adaptation at baseline velocities below as well as above the steady-state velocity optimum (Fig. 3A, B). This finding indicates that the hypothetical adaptation-induced shift of the neuronal velocity tuning towards higher values cannot entirely account for the enhancement of response transients to velocity discontinuities. In accordance with this result, consistent shifts of velocity tuning were neither found after adaptation with constant-velocity motion (Harris et al. 1999) nor after random velocity modulations (Kalb et al. 2008b). Next, it was tested whether a similar enhancement of response transients with motion adaptation is also present when the discontinuity in the visual motion stimulus is not produced by a change in velocity, but in one of the other stimulus parameters. Intriguingly, it was found that responses to discontinuities in motion direction (Fig. 3C) as well as in the contrast or the wavelength of a moving grating were also accentuated in the course of motion adaptation (Kurtz et al. 2009b). These findings suggest that adaptation improves neuronal sensitivity to sudden changes in any of the parameters of the motion stimulus.

6. Stimulus-specific adaptation as a possible explanation for improved sensitivity to stimulus discontinuities

A phenomenon, stimulus-specific adaptation has been proposed to explain why many neurons are able to maintain their sensitivity to changes in the stimulus while their responses during sustained uniform stimulation are strongly attenuated (Ulanovsky et al. 2003). Stimulus-specific adaptation provides a concept that might explain an increased sensitivity to stimulus discontinuities by adaptation without assuming a genuine shift in the stimulus-response function. Consider a neuron that integrates inputs from several elements that differ in their response optima (Fig. 2C). Adaptation with sustained stimulation activates most strongly those input elements which have their optimum close to the adapting stimulus. However, due to the shallow slope of their response function close to the peak, these inputs would show only small changes in activity in response to a stimulus discontinuity, i. e. to transient increases or decreases in stimulus intensity. A simple activity-dependent attenuation of these highly activated inputs would then improve the relative sensitivity to stimulus discontinuities. This is the case because after attenuation of strongly activated inputs the inputs that are only moderately activated by the adapting stimulus would have a relatively stronger impact on the postsynaptic integrating neuron. The overall response to stimulus discontinuities would then be enhanced relative to the background response, because the moderately activated inputs, operating in a high-slope regime of their stimulus-response function, show strong responses to stimulus changes (Fig. 2D).

Stimulus-specific adaptation has been implicated in processes that enable neurons to function as efficient 'novelty detectors', with the ability to extract the appearance of new stimulus features during sustained stimulation. For example, in the auditory system, frequency-specific adaptation is thought to facilitate the detection of novel sounds in the presence of a sustained stimulation by sound mixtures (Ulanovsky et al. 2003; Reches and Gutfreund 2008). The ability of a neuronal system to adapt in a stimulus-specific manner depends on the distinct representation of different stimulus qualities or features by individual input elements. In visual motion processing this important prerequisite for stimulus-specific adapta-

tion is at least partially met. In particular, many TCs receive input from neurons that differ in their preferred motion directions (Krapp et al. 1998; Spalthoff et al. 2010). If these inputs are attenuated during ongoing stimulation with motion, it is likely that this attenuation is strongest in those inputs that most closely match their preferred direction to the adapting motion direction. Analogous to the scheme in Fig. 2C, D, specific attenuation of strongly activated inputs could improve the sensitivity to changes in motion direction. Although it is intriguing to propose stimulus-specific adaptation as a putative mechanism underlying the enhanced sensitivity for changes in motion direction (Fig. 3C), the corresponding changes in the activity of inputs are still to be demonstrated directly in fly TCs.

Is stimulus-specific adaptation also a plausible explanation for increased sensitivity of fly TCs to discontinuities in motion velocity (Fig. 3A, B)? This form of adaptation would require inputs that differ in their velocity optimum, and are thus differentially activated by a given baseline velocity and, consequently, attenuated to different degrees during adaptation. As systematic recordings from the inputs of fly TCs are hard to obtain, differences in velocity tuning of inputs can only be indirectly deduced from the existence of multiple peaks in the velocity tuning of TCs. Whereas the velocity tuning of fly TCs appears to be smooth, distinct peaks at different velocities were shown to be present in the velocity tuning of motion-sensitive neurons of other insect species (O'Carroll et al. 1996; O'Carroll et al. 1997). Apart from enabling stimulus-specific adaptation, the coexistence of EMDs with different velocity tuning in the visual system would present a valuable prerequisite for extracting velocity independent of spatial pattern properties (Srinivasan et al. 1999).

7. Interactions between different components of motion adaptation

In the previous section, the issue of how motion adaptation changes the sensitivity to modulations in one of the stimulus parameters was addressed. However, the effects of motion adaptation on the signaling of different stimulus parameters interact with one another. For example, as will be outlined below, pronounced changes in contrast sensitivity are induced by motion adaptation, and these may in turn have a strong impact on direction sensitivity.

Harris et al. (2000) analyzed in detail how motion adaptation alters contrast sensitivity of one class of TCs, Horizontal-System (HS) neurons of the hoverfly *Eristalis tenax*. Unlike the velocity-response function or the direction tuning, which form bell-shaped or sinusoidal curves, respectively, the contrast-response function forms a sigmoid. Thus, contrast coding would profit over the entire range of contrasts from an adaptive shift of the curve's region of highest slope towards the mean contrast level of the present stimuli (as shown in Fig. 2B). A strong shift of the contrast-response function towards higher contrasts was indeed found to be present after adapting the neurons with motion of a high-contrast grating (Fig. 4). This rightward shift was not present when adapting and test stimuli were presented in different, non-overlapping regions of the large receptive field of HS-neurons (Nordström and O'Carroll 2009). This property implies that the underlying cellular adaptation process is not generated in TCs themselves, but at a more peripheral location.

In addition to the rightward shift of the contrast-response function, two further adaptation components contributed to the attenuation of contrast sensitivity: a subtractive shift of the contrast–response function

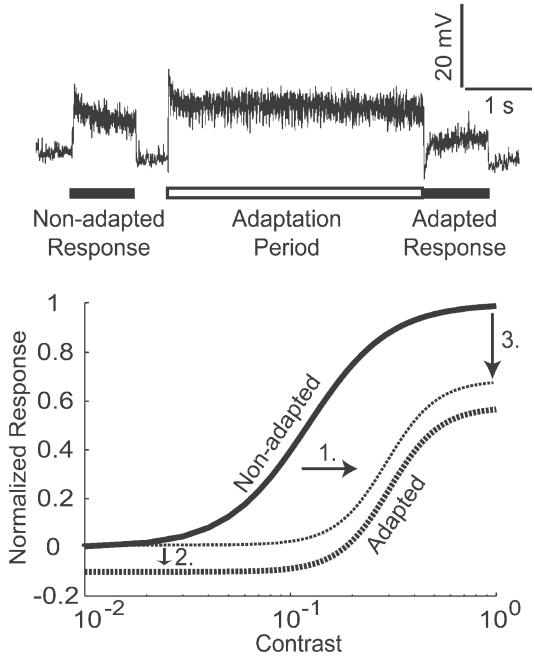


Fig. 4 Motion adaptation affects contrast gain. *Top*, response of an HS neuron of the hoverfly *Eristalis tenax* to a drifting grating of medium luminance contrast measured before and after adaptation with motion of a high-contrast grating. *Bottom*, schematic of contrast-response curves obtained by measuring non-adapted and adapted responses to drifting gratings of various contrasts. Three components of motion adaptation contribute to the attenuation of contrast sensitivity: 1) lateral shift towards higher contrast values; 2) downward shift, equivalent to neuronal after-hyperpolarization. 3) compression of the output range, which is best visible when the adapted curve is corrected for the after-hyperpolarization (thin dotted line). Modified from Harris et al. (2000)

and a compression of the output range of the neuron (Fig. 4). Whereas the origin of the latter phenomenon is unknown, the subtractive shift is caused by a depolarization-activated conductance, which manifests itself as a hyperpolarization after stimulus offset (see also Fig. 1C). This after-hyperpolarization might form the neural substrate of the famous "waterfall illusion" derived from human psychophysics, in which prolonged viewing of a strong motion stimulus induces the strong impression of motion in the opposite direction when a stationary pattern is fixated afterwards (review: Anstis et al. 1998). Interestingly, a behavioral correlate of the "waterfall illusion" has been observed in the fly's optomotor turning response (Srinivasan and Dvorak 1979). The cellular mechanism underlying the after-hyperpolarization is a depolarization-activated conductance of the TCs (Kurtz 2007; Kurtz et al. 2009a).

All the adaptation components mentioned above affect the responses of the neuron to subsequently presented motion independent of its direction. Intriguingly, adaptation of contrast sensitivity might not only modify contrast coding itself, but also direction selectivity. Changes in direction selectivity with adaptation were demonstrated in the V1-neuron of *Calliphora* (Kalb et al. 2008a). This TC prefers vertical motion, but also responds weakly to horizontal motion. Following adaptation with either vertical or horizontal motion, the ratio of vertical motion responses versus horizontal motion responses was increased in V1. A parsimonious explanation for this finding is that the responses to weak stimuli, for example, motion in a direction that differs from the neuron's preferred direction, are more affected by a subtractive shift of the contrast-response function than responses to strong stimuli.

8. Functional significance of adaptation in visual motion processing

As detailed above, motion adaptation affects directional tuning as well as contrast sensitivity of fly TCs, and it leads to enhanced

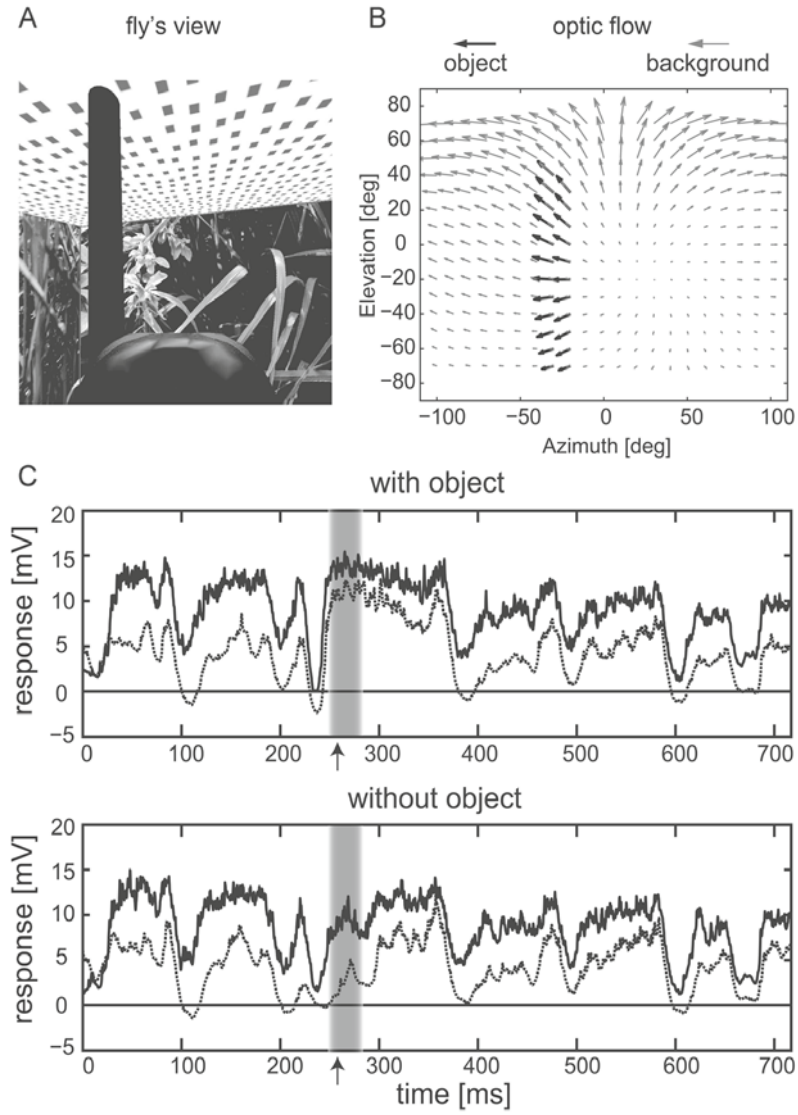


Fig. 5 Motion adaptation accentuates responses to objects in a three-dimensional environment. **A** Virtual environment presented to the fly (part of the head sketched at the bottom) in neuronal recording experiments. Trajectories and head orientations of *Calliphora* during flights in a cubic arena with patterned walls were monitored. Visual stimuli encountered by the fly were calculated and replayed on a high-speed panoramic visual stimulator during electrical recording of HS-neurons. Responses to the original image sequence were compared with those in which a virtual object (dark bar in the left half of the image) was placed close to the fly's trajectory. **B** Visual motion experienced by the fly at the instant of time shown in *A*. Arrows represent velocity vectors at different points in visual space. The object induces distinct motion cues during forward translation of the fly because it is closer to the fly than the background. **C** Average neuronal responses of 10 HS-neurons in non-adapted (solid line) and adapted state (dotted line) to image sequences with object (top) and without object (bottom). Grey shading indicates an interval during which the object is in the receptive field of the HS neuron (and the corresponding interval in the "without object" condition). Arrows indicate moment shown in **A**. In general, the neuronal response is attenuated with adaptation. However, the attenuation is much weaker when the object is in the receptive field of the neuron (upper plot, shaded area). Modified from Liang et al. (2008)

sensitivity for various types of stimulus discontinuities (Fig. 3). Adaptation might thus facilitate the extraction of those features of the stimulus that are relevant in a given behavioral situation. This view is supported by a recent study, in which *Calliphora* HS-cells were stimulated with naturalistic visual motion (Fig. 5A, B; Liang et al. 2008). A strong overall reduction in the neuronal response was observed after persistent stimulation. However, this adaptation-induced attenuation was weaker in time segments during which a virtual object moved into the receptive field (Fig. 5C). Thus, motion adaptation appears to enhance the ability of the neuron to contribute to the detection of novel signals (exemplified by the virtual object), which are potentially more important than continuing, unchanged signals.

A further benefit of adaptation may result directly from the strong decrease in overall activity and the associated decrease in energy demand. Local energy availability in a brain area may be a constraint, because neuronal signaling involves metabolically costly processes such as action potential propagation and synaptic transmission (Laughlin 2001). It has been shown that, consistent with this idea, H1 decreases its spike rate during adaptation by naturalistic stimulation (Heitwerth et al. 2005). Importantly, the information content of H1 spike trains was not reduced proportionally, resulting in considerably more information per spike.

Conclusions and outlook

At the current level of knowledge there appears to be a discrepancy between the successful phenomenological investigation of adaptation in visual motion processing neurons and the assessment of the functional significance of these phenomena for visually guided motor control in natural conditions. On the one hand, the number of different adaptation phenomena in fly motion vision described over the past 25 years is so large that only a few of them could be considered in the present chapter. On the other hand, systematic investigations of the specific roles of different components of motion adaptation in the processing of natural stimuli are still lacking.

Putative functional implications of motion adaptation, such as enhanced object detection (Liang et al. 2008), depend on how efficiently different cellular adaptation processes are activated by the complex temporal profile of a natural visual motion stimulus. Since the visual input is shaped by the fly's own movements, the resulting visual motion stimulus can have dramatically diverse dynamic characteristics (Kern et al. 2001; Boeddeker et al. 2003; Kern et al. 2005). Thus it is plausible to assume that an efficient extraction of behaviorally relevant environmental features depends on whether adaption mechanisms and its time constants are adjusted to the temporal statistics of input signals. It has been concluded that adaptation of fly motion-sensitive neurons to the statistics of random velocity modulations of a grating pattern improves the extraction of velocity information (Brenner et al. 2000). One major future task is to analyze the dynamics of visual motion occurring in natural flight situations and to test how far these temporal characteristics are matched by the dynamic properties of adaptation.

Acknowledgements

The author's labwork was supported by the Deutsche Forschungsgemeinschaft (DFG). C. Spalthoff contributed to the design of figure 1.

References

Anstis S, Verstraten FA, Mather G (1998) The motion aftereffect. Trends Cogn Sci 2: 111–117

Benda J, Herz AV (2003) A universal model for spike-frequency adaptation. Neural Comput 15: 2523–2564

Boeddeker N, Kern R, Egelhaaf M (2003) Chasing a dummy target: smooth pursuit and velocity control in male blowflies. Proc R Soc Lond B Biol Sci 270: 393–399

Borst A (2009) *Drosophila*'s view on insect vision. Curr Biol 19: R36–R47

Borst A, Egelhaaf M (1989) Principles of visual motion detection. Trends Neurosci 12: 297–306

Borst A, Flanagin VL, Sompolinsky H (2005) Adaptation without parameter change: Dynamic gain control in motion detection. Proc Natl Acad Sci U S A 102: 6172–6176

Borst A, Haag J (2002) Neural networks in the cockpit of the fly. J Comp Physiol [A] 188: 419–437

Brenner N, Bialek W, de Ruyter van Steveninck RR (2000) Adaptive rescaling maximizes information transmission. Neuron 26: 695–702

Clifford CW, Ibbotson MR, Langley K (1997) An adaptive Reichardt detector model of motion adaptation in insects and mammals. Vis Neurosci 14: 741–749

Clifford CW, Langley K (1996) Psychophysics of motion adaptation parallels insect electrophysiology. Curr Biol 6: 1340–1342

de Ruyter van Steveninck RR, Zaagman WH, Mastebroeck HAK (1986) Adaptation of transient responses of a movement-sensitive neuron in the visual system of the blowfly *Calliphora erythrocephala*. Biol Cybern 54: 223–236

Eatock RA (2000) Adaptation in hair cells. Annu Rev Neurosci 23: 285–314

Egelhaaf M (2008) Fly vision: neural mechanisms of motion computation. Curr Biol 18: R339-R341

Egelhaaf M, Borst A (1989) Transient and steady-state response properties of movement detectors. J Opt Soc Am A 6: 116–127

Egelhaaf M, Borst A, Reichardt W (1989) Computational structure of a biological motion-detection system as revealed by local detector analysis in the fly's nervous system. J Opt Soc Am A 6: 1070–1087

Egelhaaf M, Grewe J, Karmeier K, Kern R, Kurtz R, Warzecha AK (2005) Novel approaches to visual information processing in insects: case studies on neuronal computations in the blowfly. In: Christensen TA (ed) Methods in insect sensory neuroscience. CRC Press, Boca Raton, pp 185–212

Egelhaaf M, Reichardt W (1987) Dynamic response properties of movement detectors: theoretical analysis and electrophysiological investigation in the visual system of the fly. Biol Cybern 56: 69–87

Fairhall AL, Lewen GD, Bialek W, de Ruyter van Steveninck RR (2001) Efficiency and ambiguity in an adaptive neural code. Nature 412: 787–792

Fettiplace R, Ricci AJ (2003) Adaptation in auditory hair cells. Curr Opin Neurobiol 13: 446–451

Harris RA, O'Carroll DC, Laughlin SB (1999) Adaptation and the temporal delay filter of fly motion detectors. Vision Res 39: 2603–2613

Harris RA, O'Carroll DC, Laughlin SB (2000) Contrast gain reduction in fly motion adaptation. Neuron 28: 595–606

Hausen K, Egelhaaf M. (1989). Neural mechanisms of visual course control in insects. In: Stavenga DG, Hardie RC (eds) Facets of vision. Springer, Berlin, Heidelberg, pp 391–424.

Heitwerth J, Kern R, van Hateren JH, Egelhaaf M (2005) Motion adaptation leads to parsimonious encoding of natural optic flow by blowfly motion vision system. J Neurophysiol 94: 1761–1769

Kalb J, Egelhaaf M, Kurtz R (2008a) Adaptation changes directional sensitivity in a visual motion-sensitive neuron of the fly. Vision Res 48: 1735–1742

Kalb J, Egelhaaf M, Kurtz R (2008b) Adaptation of velocity encoding in synaptically coupled neurons in the fly visual system. J Neurosci 28: 9183–9193

Kern R, Lutterklas M, Petereit C, Lindemann JP, Egelhaaf M (2001) Neuronal processing of behaviourally generated optic flow: experiments and model simulations. Network: Comput Neural Syst 12: 351–369

Kern R, van Hateren JH, Michaelis C, Lindemann JP, Egelhaaf M (2005) Function of a fly motion-sensitive neuron matches eye movements during free flight. PLoS Biol 3: e171

Kohn A (2007) Visual adaptation: physiology, mechanisms, and functional benefits. J Neurophysiol 97: 3155–3164

Krapp HG (2000) Neuronal matched filters for optic flow processing in flying insects. Int Rev Neurobiol 44: 93–120

Krapp HG, Hengstenberg B, Hengstenberg R (1998) Dendritic structure and receptive-field organization of optic flow processing interneurons in the fly. J Neurophysiol 79: 1902–1917

Krapp HG, Hengstenberg R (1996) Estimation of self-motion by optic flow processing in single visual interneurons. Nature 384: 463–466

Kurtz R (2007) Direction-selective adaptation in fly visual motion-sensitive neurons is generated by an intrinsic conductance-based mechanism. Neuroscience 146: 573–583

Kurtz R, Beckers U, Hundsdorfer B, Egelhaaf M (2009a) Mechanisms of after-hyperpolarization following activation of fly visual motion-sensitive neurons. Eur J Neurosci 30: 567–577

Kurtz R, Dürr V, Egelhaaf M (2000) Dendritic calcium accumulation associated with direction-selective adaptation in visual motion-sensitive neurons in vivo. J Neurophysiol 84: 1914–1923

Kurtz R, Egelhaaf M, Meyer HG, Kern R (2009b) Adaptation accentuates responses of fly motion-sensitive visual neurons to sudden stimulus changes. Proc Biol Sci 276: 3711–3719

Kurtz R, Kalb J, Spalthoff C (2008) Examination of fly motion vision by functional fluorescence techniques. Front Biosci 13: 3009–3021

Lappe M (2000) Computational mechanisms for optic flow analysis in primate cortex. Int Rev Neurobiol 44: 235–268

Laughlin SB (2001) Energy as a constraint on the coding and processing of sensory information. Curr Opin Neurobiol 11: 475–480

Liang P, Kern R, Egelhaaf M (2008) Motion adaptation enhances object-induced neural activity in three-dimensional virtual environment. J Neurosci 28: 11 328–11 332

Lindemann JP, Kern R, van Hateren JH, Ritter H, Egelhaaf M (2005) On the computations analyzing natural optic flow: quantitative model analysis of the blowfly motion vision pathway. J Neurosci 25: 6435–6448

Maddess T, Laughlin SB (1985) Adaptation of the motion-sensitive neuron H1 is generated locally and governed by contrast frequency. Proc R Soc Lond B Biol Sci 228: 251–275

Matthews HR, Reisert J (2003) Calcium, the two-faced messenger of olfactory transduction and adaptation. Curr Opin Neurobiol 13: 469–475

Menini A (1999) Calcium signalling and regulation in olfactory neurons. Curr Opin Neurobiol 9: 419–426

Neri P, Laughlin SB (2005) Global versus local adaptation in fly motion-sensitive neurons. Proc Biol Sci 272: 2243–2249

Neri P (2007) Fast-scale adaptive changes of directional tuning in fly tangential cells are explained by a static nonlinearity. J Exp Biol 210: 3199–3208

Nordström K, O'Carroll DC (2009) The motion aftereffect: local and global contributions to contrast sensitivity. Proc Biol Sci 276: 1545–1554

O'Carroll DC, Bidwell NJ, Laughlin SB (1996) Insect motion detectors matched to visual ecology. Nature 382: 63–66

O'Carroll DC, Laughlin SB, Bidwell NJ, Harris RA (1997) Spatio-temporal properties of motion detectors matched to low image velocities in hovering insects. Vision Res 37: 3427–3439

Reches A, Gutfreund Y (2008) Stimulus-specific adaptations in the gaze control system of the barn owl. J Neurosci 28: 1523–1533

Reisenman C, Haag J, Borst A (2003) Adaptation of response transients in fly motion vision. I: Experiments. Vision Res 43: 1293–1309

Spalthoff C, Egelhaaf M, Tinnefeld P, Kurtz R (2010) Localized direction selective responses in the dendrites of visual interneurons of the fly. BMC Biol 8: 36

Srinivasan MV, Dvorak DR (1979) The waterfall illusion in an insect visual system. Vision Res 19: 1435–1437

Srinivasan MV, Laughlin SB, Dubs A (1982) Predictive coding: a fresh view of inhibition in the retina. Proc R Soc Lond B Biol Sci 216: 427–459

Srinivasan MV, Poteser M, Kral K (1999) Motion detection in insect orientation and navigation. Vision Res 39: 2749–2766

Ulanovsky N, Las L, Nelken I (2003) Processing of low-probability sounds by cortical neurons. Nat Neurosci 6: 391–398

Wark B, Lundstrom BN, Fairhall A (2007) Sensory adaptation. Curr Opin Neurobiol 17: 423–429

Webb B (2002) Robots in invertebrate neuroscience. Nature 417: 359–363

Visual motion sensing and flight path control in flies

9

Fritz-Olaf Lehmann, Peter Schützner, Hao Wang

Contents

Abstract

Flight in insects results from a feedback control loop that turns sensory information at high speed into locomotor commands. Although vision-mediated flight is considered a key factor for understanding navigation and guidance, there is a continuing debate about the exact role of the insect's compound eye for stability and maneuverability. This chapter focuses on visual motion detection in flies, highlighting the way in which these insects cope with perturbations of their visual environment. Numerical models are used to predict the precision of the fly's motor system required for heading stability from the dynamic properties of the sensory organs. Such information allows a better understanding of sensorimotor control strategies in flying insects, and is also of interest for engineers aiming to improve the performance of future generation biomimetic micro air vehicles based on nature-inspired control algorithms.

Fritz-Olaf Lehmann
University of Ulm, Institute of Neurobiology
Albert-Einstein-Allee 11, 89081 Ulm, Germany
e-mail: fritz.lehmann@uni-ulm.de

1. Introduction

Insects possess a remarkable repertoire of sophisticated aerobatic behavior such as obstacle avoidance reactions, escape responses and elaborate starting and landing programs. These maneuvers result from the interplay between genetically predetermined behavioral programs, feedback from the insect's sensory structures and mechanical forces acting on body and wings. In particular, the interaction between sensory stimuli and the formation of muscle commands for motor control has been thoroughly analyzed at different levels of investigation. Electrical recordings from visual motion-sensitive neurons such as lobula plate tangential cells (Hausen and Egelhaaf 1989; Haag et al. 1992; Egelhaaf and Borst 1993; Krapp 2000; Kern et al. 2005; Maimon et al. 2010; Rosner et al. 2010), mechano-sensory organs (Nalbach 1994; Nalbach and Hengstenberg 1994; Hengstenberg 1998; Frye 2007), and descending neurons (Gronenberg and Straus-

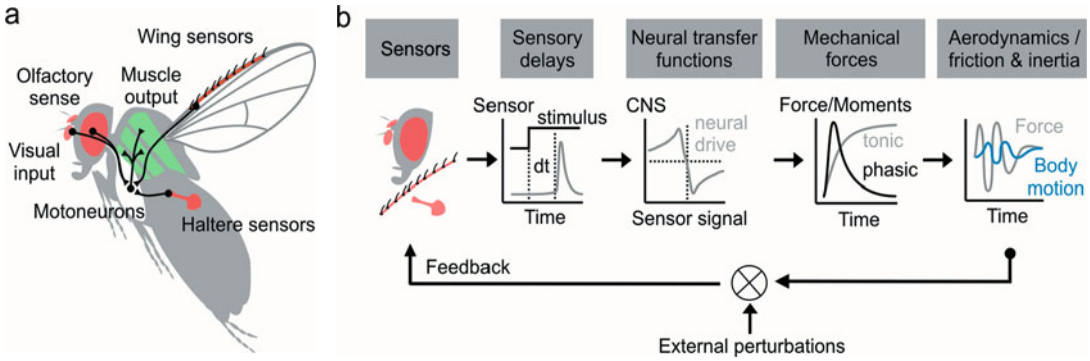


Fig. 1 **a** Sensory pathways in the fruit fly. Sensory information is projected onto motoneurons controlling two types of flight muscles: the indirect flight (power) muscles (IFM) and a set of 17 small flight control muscles at each wing. **b** Schematics of the sensory feedback control loop for flight. External perturbations mainly include displacements of the insect body due to wind and movements of the visual environment

feld 1990, 1992) were typically conducted in resting animals, whereas flight muscle activity in the fruitfly *Drosophila* (Götz 1987; Heide and Götz 1996; Lehmann and Götz 1996) and the blowfly *Calliphora* (Heide 1971; Tu and Dickinson 1996; Balint and Dickinson 2001), wing kinematics (*Drosophila*, Zanker 1990; Lehmann and Dickinson 1998), aerodynamic forces and rotational moments (*Drosophila* and house fly *Musca domestica*, Götz 1968; Götz and Wandel 1984; Lehmann and Dickinson 1998) are typically measured in tethered flying insects. Although these studies provided insights into how sensory information is processed by the central nervous system, they have so far not considered the sensorimotor feedback loop in a more *natural* context, under closed-loop feedback conditions between the insect's sensory and motor system in free flight (Fig. 1). Evaluations of the link between physical (aerodynamics) and biological (sensorimotor system) aspects of flight thus remain one of the major open challenges in insect flight research.

Vision in insects serves three main functions: First, it provides cues for navigation, orientation and guidance, as shown, for example, in the foraging behavior of the honey bee (Dyer and Gould 1981; Gould 1986; Srinivasan et al. 2000; Si et al. 2003), migration flights of butterflies (Brown and Chippendale 1974; Dudley 1991) and the fast maneuvers in flies (*Drosophila*, Tammero and Dickinson 2002a, 2002b; Fry et al. 2003). Second, it provides information about the structure of the visual environment and helps detection of attractive or abhorrent objects (Reichardt et al. 1989; O'Carrol 1993; Zanker 1993), and third, it generates feedback signals for stabilization of the insect's head and body (Heisenberg and Wolf 1984). The latter function is of particular interest, given the limited encoding capacity of biological sensors. For some insects, it has been suggested that control of body orientation is achieved mainly by mechano-sensory feedback coming from gyroscopic organs that vibrate at wing stroke frequency. Evolutionary transformed hindwings, called halteres, encode angular velocities around the three body axes in flies (*Drosophila*, Dickinson et al. 2001; Frye 2007; *Calliphora*, Nalbach 1994; Hengstenberg 1998), while in the functionally two-winged tobacco hornworm moth *Manduca sexta*, head rotations are sensed by the Coriolis forces induced in the large, vibrating antennae. If these sensory structures are cut, the animals are not able to achieve stable flight

(Sane et al. 2007). Although the above findings highlight the importance of mechanical feedback, they do not tell us much about how vision and mechano-sensory feedback signals are integrated by the brain in an *intact* animal. Even worse, recent studies on the aerodynamics of animal flight have emphasized that turning movements are stabilized *passively* by means of high frictional wing damping, suggesting only a minor need for sensory control (Mayer et al. 1988; Hesselberg and Lehmann 2007; Ramamurti and Sandberg 2007; Hedrick et al. 2009). The roll damping coefficient in birds, for example, is 2–6 times the coefficients typical of airplane flight dynamics, which limits roll magnitude during maneuvering flight (Hedrick 2007; Hedrick et al. 2007). Similar results were also found for damping of yaw turning in the fruit fly (Hesselberg and Lehmann 2007), the tobacco hornworm moth, hummingbirds, and fruit bats (Hedrick et al. 2009). These animals cover a 10-decade range of inertial moments during rapid yaw turning, indicating the significance of passive damping for animal flight *per se*. Since passive damping helps to bring down the stimulus bandwidth into a range the visual and mechano-sensory system can encode, it is possible that our current view overstates the neuronal requirements needed for stability and maneuverability in flying insects.

In contrast to *in vitro* approaches, it is difficult to measure the output of a sensory organ in a freely maneuvering insect and under natural sensory feedback conditions. A way to circumvent this problem is to conduct replay experiments in which the output of the sensory system is numerically modeled according to the sensory input and the sensor's response characteristics (transfer function, Fig. 1). A famous example of this concept is found in reconstructions of the output generated by the elementary motion detection system in freely cruising flies, in which optic flow on the fly's retina is reconstructed by high-speed video motion analysis of animals flying freely in a visual scenery (see chapter II,6 by R Kurtz; Reichardt and Poggio 1975, 1976). The meaning of simulated outputs in replay experiments, however, is limited and ignores efference copies and other physiological changes of neurons and muscles occurring *in vivo*. Octopamine, for example, is a common neuromodulator in insects that has a pronounced impact on the response characteristics of neurons such as lobula plate tangential cells (Longden and Krapp 2009) and on muscle performance (Chapman 1998). Fruit flies with genetically altered biogenic amine levels thus show significant reductions in flight duration, initiations and maintenance compared to wild types (Brembs et al. 2007). Moreover, sensory modeling strongly relies on the precision of the input signals reconstructed from the body movements of the freely flying animal and thus, in the case of the visual system, on the reconstructions of optic flow on the animal's retina (retinal optic flow). Given the small body size of insects, such reconstructions are challenging because the insect's gaze depends not only on the orientation of the body in space, but also on the relative position of the head with respect to the body. Tethered flying fruit flies *Drosophila* may rotate their heads around the three rotational axes up to 120°, i. e. yaw (±25°), pitch (±35°) and roll (±120°, Hengstenberg 1991). In unrestrained fruit flies, flying inside a random-dot visual arena, we measured smaller angles of approximately ±11°, ±21° and ±18°, respectively. Recordings of freely flying blowflies using Helmholtz coils affixed to the head and the thorax also suggest that the animals attempt to stabilize their retinal image by performing small head saccades while turning (van Hateren and Schilstra 1999; Kern et al. 2005). So far, however, no experimental study has simultaneously scored motor output (wing kinematics) and visual input in freely flying insects.

2. Vision-mediated flight control in flies

2.1 Optomotor behavior in tethered flight

Early studies of vision-mediated flight control demonstrated that tethered flies produce yaw, roll and pitching moments around their three body axes, in response to the rotation of a visual environment (optomotor response, Götz 1968; Reichardt and Poggio 1976; Blondeau and Heisenberg 1982; Kirchner and Srinivasan 1989). Under closed-loop conditions, optomotor behavior persists as long as the external visual stimulus produces retinal slip on the animal's compound eye (Wolf and Heisenberg 1990). Optomotor yaw turning behavior in insects is considered as operating via feedback in which an increase in neural activity produced by faster front-to-back retinal slip in one eye generates increased flight thrust on the contralateral body side of the animal. As a consequence, an animal achieves straight flight when optic flow is similar on both eyes (optomotor equilibrium, Götz 1975; for review see Wehner 1981). The optomotor reflex has been successfully implemented in robotic platforms and numerical models (Huber et al. 1999; Neumann and Bülthoff 2001; Iida 2003; Reiser and Dickinson 2003).

The conventional view that the optomotor steering response in fruit flies is driven by rotational motion cues has recently been questioned by Tammero and Dickinson (2004). The authors suggest that in the case of yaw, translational motion cues generated by laterally centered foci of expansion and contraction may fully account for turning behavior in yaw. At the level of visual motion detection this finding implies that steering responses to image rotation might emerge from a visual system organized to detect translation-induced flow fields, rather than from a rotation sensitive system. This result is also supported by an analysis of fruit flies flying freely in a stationary environment (Tammero and Dickinson 2002a, 2002b). The latter studies demonstrate that lateral expansion of visual cues may initiate a flight saccade while the asymmetry in the output of the local motion detector prior to the saccade primarily influences the direction, but not the turning angle, of the saccade. There is also evidence that in a stationary environment freely flying flies gradually turn away from the side experiencing a stronger motion stimulus, a response opposite to that predicted from a conventional model based upon optomotor equilibrium. This phenomenon was first demonstrated in bees (Kirchner and Srinivasan 1989).

2.2 Optomotor behavior in free flight

The fundamental concepts about visuomotor control mechanisms in flies gained from tethered flight studies have also been tested in free flight experiments on various fly species. Compared to insects that are tethered inside a flight simulator, freely flying animals may employ various strategies to achieve a reduction of retinal flow on both compound eyes and thus attain optomotor equilibrium. In the following section, we summarize behavioral responses scored in flies facing perturbations of the visual environment, i.e. rotation of a panorama, during free flight (Fig. 2a).

Side slip maneuvers in which the animal orients its body normally to the visual environment may reduce retinal slip in the frontal region of the animal's visual field, while producing translational optic flow fields with lateral foci of expansion and contraction on either eye, respectively. Thus this behavior cannot totally compensate for retinal slip caused by the rotational visual stimulus. Side slipping was demonstrated in male hoverflies *Syritta pipiens* (Collett

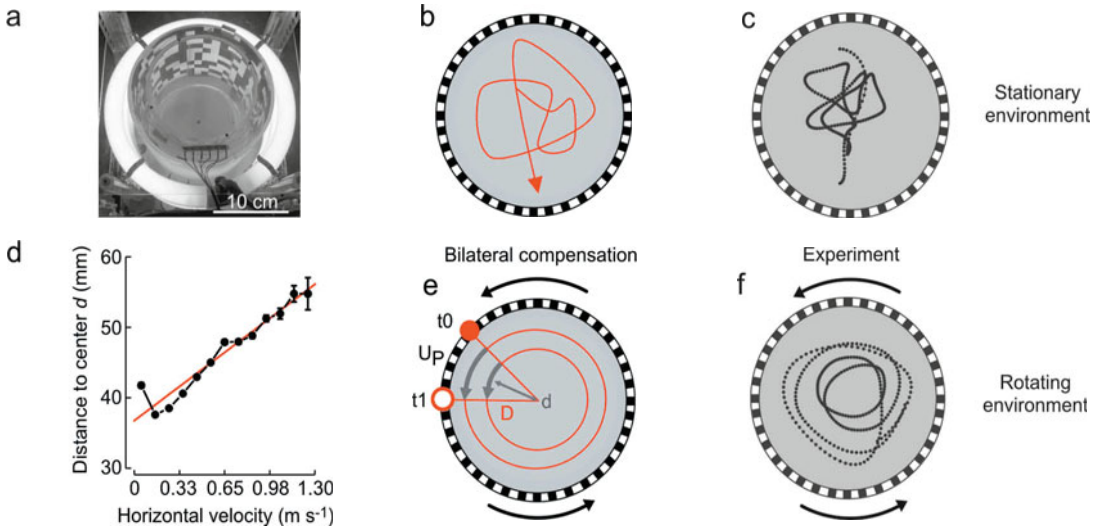


Fig. 2 Behavioral responses of fruit flies to perturbations of the visual environment. **a** Free flight arena (14 cm diameter, 25 cm height) that allows rotation of a random dot visual pattern. **b**, **e** Examples of hypothetical flight paths for a stationary (upper row) and rotating visual environment (lower row) assuming self-induced visual flow in **b** and complete compensation of retinal slip on each of the eyes in **e** (red, concentric flight trajectories). d: distance between fly and arena center, D: distance between pattern and arena center, U_P: translational (linear) velocity of pattern. At complete slip compensation, the animal matches both its yaw turning rate to the angular velocity of the rotating pattern and its forward speed to the product between U_p, d and D^{-1}. **c**, **f** Flight path of a freely cruising fruit fly in a stationary and rotating (500° s^{-1}, counter clockwise) environment. **d** Measured increase in binned heading velocity confirms predictions by the model shown in **e**. *N* = 131 flies, means ±S. D. Data sampling interval: 8 ms, red line indicates linear regression fit to data

1980a, 1980b). Hoverflies also employ a second strategy for retinal slip compensation because their locomotor system allows them to rotate around the yaw axis without extensively translating sideways, forward or backward: they turn around their vertical axes while hovering with low forward speed at the *centre* of the surrounding panorama (Fig. 2). This behavior allows the animal to completely compensate for rotational visual perturbations. In hoverflies, retinal slip compensation via side slip flight, however, occurs more often than via yaw turning. A possible explanation for this finding is the specialized frontal eye region in male hover flies favoring smooth object tracking (Collett and Land 1975).

Fruit flies employ a behavioral strategy that involves both forward and turning flight (Mronz and Lehmann 2008). Although this species is capable of hovering flight at low forward speed, the insect responds to the visual stimuli by rotating *and* translating inside the visual panorama (Fig. 2e, f). Employing this strategy, the fly may achieve zero retinal slip on both compound eyes flying at any point of the flight arena, regardless of its distance to the centre: The only prerequisite is that turning velocity matches the angular velocity of the rotating panorama and that the fly's horizontal velocity is equal to the product of translational (linear) velocity of the panorama and the ratio between the fly's distance from the drum center and the drum radius (Fig. 2d–f). At elevated forward velocity above approximately 0.45 ms^{-1}, however, the required forces needed for turning exceed the locomotor

capacity of the animal. This is mainly due to the production of centripetal forces needed to keep the fly on track that account for up to 70 % of the fly's locomotor reserves (Mronz and Lehmann 2008). Experiments in which the angular velocity of a rotating visual panorama is low (100 to 500° s^{-1}) thus show only a small mismatch between the required (zero retinal slip) and measured translational and rotational velocities of the animal body. At rotational velocities of the visual panorama above 500° s^{-1}, by contrast, the fly is not capable of exactly matching its behavior to those velocities required for zero-slip condition. At this point the animal changes its strategy and performs straight flight interspersed by sudden and fast 120° yaw turns, termed body saccades, in the direction of the moving visual environment (Mronz and Lehmann 2008). The latter behavior was also observed in the housefly *Musca domestica*. Although *Musca* produces continuous yaw moments in response to a rotating visual panorama when flown under tethered flight conditions, the predominant free flight behavior in response to a 2.5 Hz horizontally oscillating visual panorama consists of rather straight segments and fast saccades (Wagner 1986).

3. Modeling visual motion sensing in free flight

The output signal of the Hassenstein-Reichardt elementary motion detector (EMD) of the insect compound eye may be derived from a set of equations that have been developed from both behavioral experiments on tethered insects and electrophysiological recordings from motion-sensitive neurons in the lobula plate of the fly's brain (Reichardt and Poggio 1975; Egelhaaf et al. 1989, 1993; Kern et al. 2005). The EMD was later modified according to additional data on the phasic-tonic response properties of the motion sensitive system (Fig. 3a, Kern and Egel-

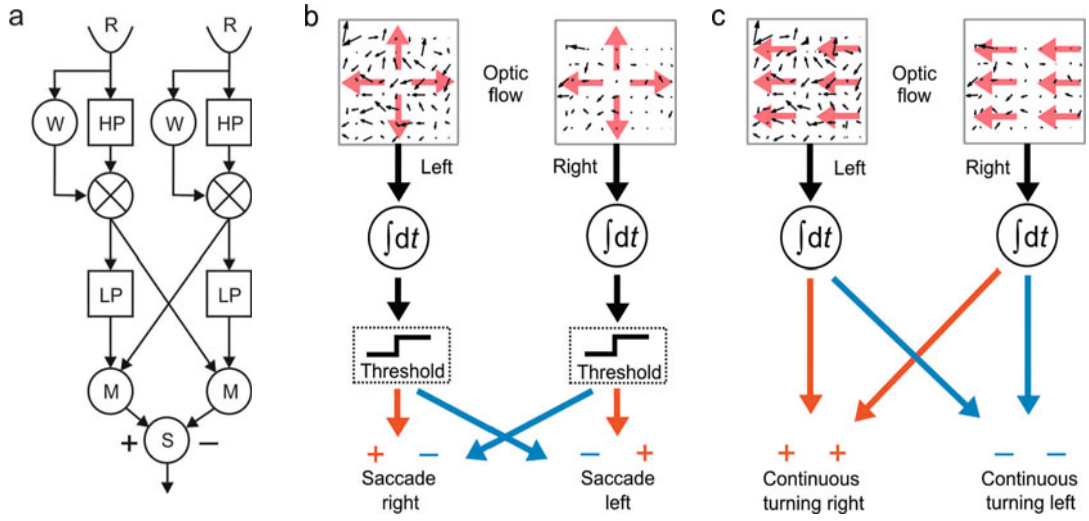


Fig. 3 **a** Schematics of the insect elementary motion detector (EMD). **b** Schematics of a visual-expansion- and **c** rotation-sensitive visual system based on optic flow of the left and right compound eye. The systems ('matched filters') provide information for controlling saccadic flight turns and smooth turning behavior, respectively. R: receptor, W: weighting function, HP: high pass filter, LP: low pass filter, M: multiplication unit, S: sum, dt: time step. **b** and **d** were modified from Tammero and Dickinson (2004)

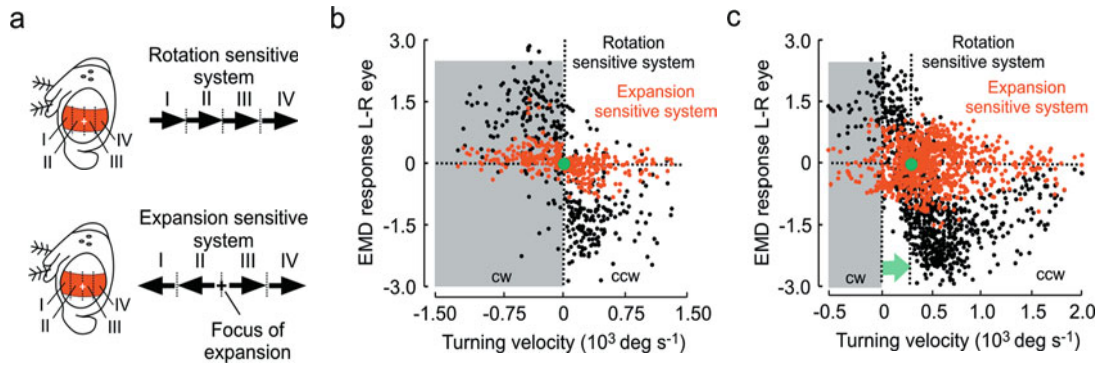


Fig. 4 **a** Outputs of a rotation- and an expansion-sensitive two-dimensional elementary motion detector (EMD) system, calculated from measurements of a fruit fly cruising freely in a **b** stationary and **c** rotating (300° s^{-1}, counter clockwise) flight arena (*cf.* Fig. 2). EMD results are plotted against the fly's yaw turning rate. Zero set point of the feedback control loop needed for retinal slip compensation is indicated by a green dot. Green arrow in **c** shows shift in set point due to visual stimulation. cw: clockwise rotation, ccw: counter clockwise rotation of the fly, L: left eye, R: right eye, flight time: 7.7 s, data sampling interval: 8 ms

haaf 2000). There are numerous articles on the function and implementation of this system including its application to insect flight. Simulations of visual motion detection in the fruit fly may be found in Götz and Buchner (1978), Tammero and Dickinson (2002 a), and Mronz and Lehmann (2008). Linear summation of EMD outputs for adjacent ommatidia in a 2-dimensional field of view allows constructions of 'matched filters', systems that are sensitive to stimuli caused by a rotating visual environment (rotation-sensitive detector system) or to expansion/contraction of visual stimuli in a lateral focus (expansion sensitive detector system) on each of the two compound eyes (Fig. 3 b and c, Krapp et al. 1998, 2001). The outputs of the matched filters may be numerically simulated by projection of the visual environment that has been experienced by the freely flying insect, onto the two model eyes. Consequently, the simulation of the two detector systems allows us to derive the driving sensory signals for the two types of yaw turning control in free flight: continuous/smooth (< 1000° s^{-1} turning rate) and saccadic (> 1000° s^{-1} turning rate) flight turns.

Data-based simulations of matched filter outputs in *Drosophila* confirm the assumption of a behavioral-mediated no-slip condition on *each* of the two compound eyes. Thus, during visual perturbations of the fly's control system by rotation of the visual environment, the average matched filter output of the rotation and translation sensitive systems is scattered around zero. Figure 4 shows an example of how the fruit fly re-establishes the set point (green) of the rotation-sensitive system when flown in a rotating environment. Optomotor balance (zero EMD response) is established exactly when the fly's turning velocity matches the angular velocity of the rotating visual environment. The expansion-sensitive system, by contrast, seems not to be linearly correlated with smooth yaw turning velocity in freely flying fruit flies. Fast-Fourier-Transformation analyses of the output of the rotation-sensitive matched filters further show that the EMD signal oscillates sinusoidally during free flight in *Drosophila*, with frequencies ranging from approximately 2 to 5 Hz (Mronz and Lehmann 2008). At present, it is still unclear whether the animal produces these oscilla-

tions by *actively* varying its yaw moments, or whether this feedback is simply due to the latency with which the eye transforms the locally changing brightness on the retina into motion information. Alternatively, such oscillations in sensory feedback may also result from the imprecision of the motor system and thus from the difficulty of the flight musculature to finely control wing kinematics and thus yaw moments during flight. In the next section, we thus highlight the link between response latencies of sensory systems and the precision of the muscle system needed for turning control.

4. The role of feedback for motor precision

The speed with which an insect detects and processes sensory information is a key factor for understanding the *limits* of the sensorimotor feedback control loop for body control. The compound eyes and the halteres, differ both in their response delays to an external stimulus and in their sensory bandwidths (linear encoding range) for angular turning rates around the three body axes. The flies' halteres are driven by their own power and flight control muscles (Nalbach 1993), beat in anti-phase with the wings and sense changes in Coriolis forces when the animal's body rotates (Pringle 1948; Nalbach 1994; Nalbach and Hengstenberg 1994). The haltere's sensory axons project to motor neurons of the wing control muscles via fast electrical synapses (Fayyazuddin and Dickinson 1996). Experiments on tethered fruit flies show that the halteres may encode rotational velocities approximately within a $\pm 700^\circ\,s^{-1}$ range and provide feedback via the electrical synapse within a single wing stroke (Sherman and Dickinson 2003). In the fruit fly, the delay of this sensory organ is thus less than 5 ms (200 Hz stroke frequency). The vision system in flies, by contrast, exhibits a smaller linear encoding range for rotational motion between approximately $\pm 125^\circ\,s^{-1}$ but may detect the direction of turning at least up to $\pm 1\,000^\circ\,s^{-1}$ (*Calliphora*, Sherman and Dickinson 2003). It is important to note that these values are derived from *in vitro steady-state* responses of the visual system, ignoring the dynamics of visual feedback during flight. Compared to the halteres, the vision system displays at least a 6-fold larger response delay to the onset of visual motion. Heisenberg and Wolf (1988) reported 100 ms collision avoidance delay in tethered flying fruit flies and David (1985) approximately 50 ms response delay in freely flying fruit flies. Cross-correlation analysis between turning rate and the output of an expansion sensitive EMD system suggests 74 ms for free flight maneuvering in the fruit fly (Mronz and Lehmann 2008). In freely flying house flies *Musca domestica*, the values are scattered around 30 ms during male-female chases (Land and Collett 1974). Electrophysiological recordings show that most of this delay is due to the photon-transduction process in the compound eye rather than to delays in encoding visual information or the activation of the flight muscle system. In flies photo-transduction latency varies between approximately 30 ms (fruit fly, Hardie and Raghu 2001) and 75 ms (house fly, Howard et al. 1984).

To evaluate the consequences of the physiological limits of sensory organs for flight stability and body control, we combine the transfer functions and response delays for sensory feedback with the physical properties of the fly body. On a physical level, an insect's instantaneous angular velocity during turning ω, at a given time t, results from the three major components: the torque T produced around the vertical or horizontal

body axes, the mass moment of inertia (rotational inertia) *I*, and the frictional damping coefficient *C* of body and wings (Fry et al. 2003; Hesselberg and Lehmann 2007). This relationship may be written as,

$$T(t) = I\dot{\omega}(t) + C\omega(t)\,. \qquad (1)$$

Moments of inertia mainly depend on the distribution of body mass around the rotational axes and thus on body shape and body posture during flight. In the fruit fly, inertial moments are small and inertial yaw moment only amounts to 0.52 pNms^2. The frictional damping coefficient, by contrast, amounts to 54 pNms for yaw maneuvers and significantly limits the maximum turning rate during flight. At mean yaw torque of 1.0 pNm for saccadic turning, the fruit fly's turning rate thus saturates rapidly within 3–4 stroke cycles or 15–20 ms at approximately 3500° s^{-1} (Hesselberg and Lehmann 2007; Ramamurti and Sandberg 2007). Besides its significance in *limiting* turning rate, wing damping is highly effective in *terminating* body rotations and thus reduces the need for active braking moments at the end of a body saccade. The benefit of *passive breaking* by high frictional damping was demonstrated in a large variety of flying animals covering flies, moths, bats and birds (Hedrick et al. 2009).

Despite the apparent benefit of aerodynamic damping for flight, the above equation implies that the animal's motor system may generate body turning rates *outside* the bandwidth of the fly's sensory systems. *Maximum torque* production is thus a vital measure for an insect in order for it to stay within the limits of the sensory apparatus during maneuvering flight. For example, a long sensor delay in conjunction with a small sensor bandwidth, requires small rotational moments and thus precise kinematic control of the flapping wings. By contrast, a short sensory delay paired with a large sensor bandwidth allows the animal to steer with larger torque pulses, without exceeding the limits of the sensory apparatus. The latter case thus reduces the need for precise motor control. The reason for the above relationships is that the sensory response delay equals the angular acceleration period within a turning maneuver. We may thus hypothesize that the precision of the flight muscle system matches, at least to some degree, the capacity (bandwidth and delay) of the sensory system in an insect. Based on these considerations, *maximum* torque T_{Max} may be estimated by replacing angular speed in equation (1) by the sensor's upper (absolute) threshold for detecting rotational body movement ω_V and angular acceleration by the ratio between the sensor's encoding bandwidth and its response delay t_V (Hesselberg and Lehmann 2007). A time invariant version of equation (1) may be then written as:

$$T_{Max} = I\omega_V t_V^{-1} + C\omega_V\,. \qquad (2)$$

If we further convert yaw torque into the bilateral differences of wing stroke amplitude using a conversion factor 2.9 × 10^{-10} Nm·deg^{-1} stroke amplitude (Götz 1983), we obtain the maximum changes in the left and right stroke amplitudes allowed to keep the sensory stimulus within the *linear encoding range* of the halteres and the vision system, respectively. In other words: The combined modeling of physical properties and sensory capabilities allows predictions of the required precision of muscle control during turning behavior and also comparisons between the two main sensory organs for flight control.

To estimate the required precision of muscle control during vision-mediated turning flight in the fruit fly, we may calculate torque-equivalent stroke amplitudes using 30 ms response delay for the photo-transduction process (Hardie and Raghu 2001). Figure 5a shows how the required precision

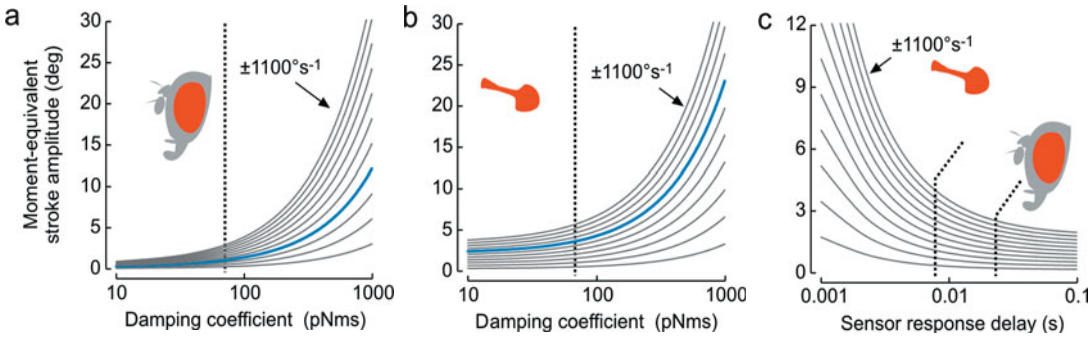


Fig. 5 Numerical simulation of motor precision for yaw turning control in the fruit fly *Drosophila*. The graphs show the *maximum* yaw-torque-equivalent left-minus-right stroke amplitude allowed to keep the sensory stimulus within the bandwidth of the vision in **a** and the haltere system in **b** [*cf.* equation (2)]. Dotted line in **a** and **b** indicates the natural frictional damping coefficient of wings and body during yaw turning (54 pNms). **a** Results for the visual and **b** the mechano-sensory system (halteres), based on 30 and 5 ms sensor response delay (dotted lines in **c**), respectively. **c** Maximum changes in stroke amplitudes plotted against sensor response delay. Grey lines were calculated assuming different bandwidths of the vision and haltere system (spacing: 100° s^{-1}). Blue indicates 400° s^{-1} in **a** and 700° s^{-1} in **b**. See text for further information

of amplitude control changes with changing frictional damping at sensory bandwidths for the EMD ranging from 100 to 1 100° s^{-1} (grey curves). As predicted above, maximum yaw torque-equivalent difference between left and right stroke amplitude for turning must decrease with both decreasing frictional damping and decreasing bandwidth of visual motion detection. The numerical model suggests that at the fruit fly's natural damping coefficient of 54 pNms (dotted line, Fig. 5a and b) and assuming ±390° s^{-1} sensory bandwidth of the vision system (Hesselberg and Lehmann 2007), the difference of stroke amplitude may not exceed a small value of 1.0 to 2.0°. Even at a much higher linear sensory bandwidth of ±1 100° s^{-1} for visual motion detection, yaw torque-equivalent stroke amplitudes should not exceed 2.6°, in order to keep body turning rate within the range of the vision system.

By contrast, the smaller response delay of 5 ms (or less) as suggested for the halteres, allows the fly to employ larger differences between left and right stroke amplitudes during yaw steering (Fig. 5b). For example, assuming the same sensory bandwidth of ±700° s^{-1} for both sensory systems, the haltere-mediated flight requires 47 % less precision in kinematic yaw torque control than the vision-mediated control (3.5° vs. 1.6° stroke amplitude difference, Fig. 5c). However, this result also suggests that the *relative* benefit of haltere-mediated yaw turning control in the fruit fly is small compared to vision-mediated control, despite the 6-fold shorter response delay of the halteres.

Conclusions

Evaluations of the relative significance of sensory feedback for guidance and stability in flying insects are complex. In contrast to previous work, recent studies show that approaches based solely on the physiological properties of muscles and neurons may not fully explain body stabilization and path control, because they ignore the fluid dynamic interactions between the animal and its surrounding air. Consequently, body stabilization and maneuverability are two sides of the same coin. High frictional damping on body and wings favors body stability and lowers the need for elaborate sensory

feedback, but at the cost of reduced agility and maneuverability. The reverse is true of insects that rely on high agility, as the flight system should be able to produce high locomotor forces in order to cope with frictional damping, paired with fast feedback and a precise muscle system. Agile insects such as flies might thus benefit from having some motor learning capacity that allows a fine tuning of their genetically preprogrammed feedback control loop. An experimental study on the role of flight experience in fruit flies for turning control recently confirmed this prediction, showing that previous free flight experience adjusts locomotor fine tuning of several key parameters for flight (Hesselberg and Lehmann 2009).

References

Balint CN, Dickinson MH (2001) The correlation between wing kinematics and steering muscle activity in the blowfly *Calliphora vicina*. J Exp Biol 204: 4213–4226

Blondeau J, Heisenberg M (1982) The three dimensional optomotor torque system of *Drosophila melanogaster.* J Comp Physiol A 145: 321–329

Brembs B, Christiansen F, Pflüger HJ, Duch C (2007) Flight initiation and maintenance deficits in flies with genetically altered biogenic amine levels. J Neurosci 27: 11122–11131

Brown JJ, Chippendale GM (1974) Migration of the monarch butterfly, *Danaus plexippus*: energy resources. J Insect Physiol 20: 1117–1130

Chapman RF (1998) The insects: structure and function. Cambridge University Press, Cambridge, UK

Collett TS (1980a) Some operating rules for the optomotor system of a hoverfly during voluntary flight. J Comp Physiol A 138: 271–282

Collett TS (1980b) Angular tracking and the optomotor response: an analysis of visual reflex interaction in a hoverfly. J Comp Physiol A 140: 145–158

Collett TS, Land MF (1975) Visual control of flight behavior in the hoverfly, *Syritta pipiens* L. J Comp Physiol A 99: 1–66

David CT (1985) Visual control of the partition of flight force between lift and thrust in free-flying *Drosophila*. Nature 313: 48–50

Dickinson MH, Birch J, Fry S et al. (2001) Deconstructing the aerodynamics of insect flight. Meeting of the Society of Integrative and Comparative Biology, 41

Dudley R (1991) Biomechanics of flight in neotropical butterflies: aerodynamics and mechanical power requirements. J Exp Biol 159: 335–357

Dyer FC, Gould JL (1981) Honey bee orientation: A backup system for cloudy days. Science 214: 1041–1042

Egelhaaf M, Borst A (1993) A look into the cockpit of the fly: visual orientation, algorithms, and identified neurons. J Neurosci 13: 4563–4574

Egelhaaf M, Borst A, Reichardt W (1989) Computational structure of a biological motion-detection system as revealed by local detector analysis in the fly's nervous system. J Opt Soc Am A 6: 1070–1087

Egelhaaf M, Borst A, Flecks S, Wildemann A (1993) Neural circuit tuning fly visual interneurons to motion of small objects. II. Input organization of inhibitory circuit elements revealed by electrophysiological and optical recording techniques. J Neurophysiol 69: 340–351

Fayyazuddin A, Dickinson MH (1996) Haltere afferents provide direct, electrotonic input to a steering motor neuron of the blowfly, *Calliphora*. J Neurosci 16: 5225–5232

Fry SN, Sayaman R, Dickinson MH (2003) The aerodynamics of free-flight maneuvers in *Drosophila*. Science 300: 495–498

Frye MA (2007) Behavioral neurobiology: A vibrating gyroscope controls fly steering maneuvers. Curr Biol 17: 134–136

Götz KG (1968) Flight control in *Drosophila* by visual perception of motion. Biol Cybernetics 4: 199–208

Götz KG (1975) Sehen, Abbilden, Erkennen – Verhaltensforschung am visuellen System der Fruchtfliege *Drosophila*. Verh Schweiz Naturforsch Ges: 10–33

Götz KG (1983) Bewegungssehen und Flugsteuerung bei der Fliege *Drosophila*. In: Nachtigall W (ed) BIONA-report 2. Fischer, Stuttgart

Götz KG (1987) Relapse to 'preprogrammed' visual flight-control in a muscular subsystem of the *Drosophila* mutant 'small optic lobes'. J Neurogenetics 4: 133–135

Götz KG, Buchner E (1978) Evidence for one-way movement detection in the visual system of *Drosophila*. Biol Cybernetics 31: 243–248

Götz KG, Wandel U (1984) Optomotor control of the force of flight in *Drosophila* and *Musca* II.

Covariance of lift and thrust in still air. Biol Cybernetics 51: 135–139

Gould JL (1986) The locale map of the honey bees: Do insects have cognitive maps? Science 232: 861–863

Gronenberg W, Strausfeld NJ (1990) Descending neurons supplying the neck and flight motor of Diptera: Physiological and anatomical characteristics. J Comp Neurol 302: 973–991

Gronenberg W, Strausfeld NJ (1992) Premotor descending neurons responding selectively to local visual stimuli in flies. J Comp Neurol 316 (1): 87–103

Haag J, Egelhaaf M, Borst A (1992) Dentritic integration of motion information in visual interneurons of the blowfly. Neurosci Letters 140: 173–176

Hardie CR, Raghu P (2001) Visual transduction in *Drosophila*. Nature 413: 186–193

Hausen K, Egelhaaf M (1989) Neural mechanisms of visual course control in insects. In: Stavenga DG, Hardie RC (eds) Facets of vision. Springer, Berlin, Heidelberg

Hedrick TL (2007) Experimental study of low speed turning flight in cockatoos and cockatiels. AIAA 2007–0044

Hedrick TL, Usherwood JR, Biewener AA (2007) Low speed maneuvering flight of the rose-breasted cockatoo (*Eolophus roseicapillus*). II. Inertial and aerodynamic reorientation. J Exp Biol 210: 1912–1924

Hedrick TL, Cheng B, Deng X (2009) Wingbeat time and the scaling of passive rotational damping in flapping flight. Science 324: 252–255

Heide G (1971) Die Funktion der nicht-fibrillären Flugmuskeln bei der Schmeißfliege *Calliphora*. Teil II: Muskuläre Mechanismen der Flugssteuerung und ihre nervöse Kontrolle. Zool Jb Physiol 76: 99–137

Heide G, Götz KG (1996) Optomotor control of course and altitude in *Drosophila* is achieved by at least three pairs of flight steering muscles. J Exp Biol 199: 1711–1726

Heisenberg M, Wolf R (1984) Vision in *Drosophila*. Springer-Verlag, Berlin

Heisenberg M, Wolf R (1988) Reafferent control of optomotor yaw torque in *Drosophila melanogaster*. J Comp Physiol A 163: 373–388

Hengstenberg R (1991) Body posture, head posture and gaze movements in the fruitfly. Verh Dtsch Zool Ges 8: 344–345

Hengstenberg R (1998) Controlling the fly's gyroscope. Nature 392: 757

Hesselberg T, Lehmann F-O (2007) Turning behaviour depends on frictional damping in the fruit fly *Drosophila*. J Exp Biol 210: 4319–4334

Hesselberg T, Lehmann F-O (2009) The role of experience in flight behaviour of *Drosophila*. J Exp Biol 212: 3377–3386

Howard J, Dubs A, Payne R (1984) The dynamics of phototransduction in insects: A comparative study. J Comp Physiol A 154: 707–718

Huber SA, Franz MO, Bülthoff HH (1999) On robots and flies: modelling the visual orientation behaviour of flies. Rob Auton Syst 29: 227–242

Iida F (2003) Biologically inspired visual odometer for navigation in a flying robot. Rob Auton Syst 44: 201–208

Kern R, Egelhaaf M (2000) Optomotor course control in flies with largely asymmetric visual input. J Comp Physiol A 186: 45–55

Kern R, van Hateren JH, Lindemann JP, Egelhaaf M (2005) Function of a fly motion-sensitive neuron matches eye movements during free flight. PLoS Biology 3: 1130–1138

Kirchner W, Srinivasan MV (1989) Freely-flying honeybees use image motion to estimate object distance. Naturwissenschaften 76: 281–282

Krapp H (2000) Neuronal matched filters for optic flow processing in flying insects. In: Lappe M (ed) Neuronal processing of optic flow. Academic Press, San Diego, San Francisco, New York

Krapp HG, Hengstenberg B, Hengstenberg R (1998) Dentritic structure and receptive-field organization of optic flow processing interneurons in the fly. J Neurophysiol 79(4): 1902–1917

Krapp HG, Hengstenberg R, Egelhaaf M (2001) Binocular contributions to optic flow processing in the fly visual system. J Neurophysiol 85(2): 724–734

Land MF, Collett TS (1974) Chasing Behaviour of houseflies (*Fannia canicularis*). J Comp Physiol 89: 331–357

Lehmann F-O, Götz KG (1996) Activation phase ensures kinematic efficacy in flight-steering muscles of *Drosophila melanogaster*. J Comp Physiol A 179: 311–322

Lehmann F-O, Dickinson MH (1998) The control of wing kinematics and flight forces in fruit flies (*Drosophila* spp.). J Exp Biol 201: 385–401

Longden KD, Krapp HG (2009) State-dependent performance of optic-flow processing interneurons. J Neurophysiol 102: 3606–3618

Maimon G, Straw AD, Dickinson MH (2010) Active flight increases the gain of visual motion

processing in *Drosophila*. Nature Neurosci 13: 393–401

Mayer M, Vogtmann K, Bausenwein B, Wolf R, Heisenberg M (1988) Flight control during 'free yaw turns' in *Drosophila melanogaster*. J Comp Physiol A 163: 389–399

Mronz M, Lehmann F-O (2008) The free flight response of *Drosophila* to motion of the visual environment. J Exp Biol 211: 2026–2045

Nalbach G (1993) The halteres of the blowfly *Calliphora* I. kinematics and dynamics. J Comp Physiol A 173: 293–300

Nalbach G (1994) Extremely non-orthogonal axes in a sense organ for rotation: behavioral analysis of the dipteran haltere system. Neurosci 61: 149–163

Nalbach G, Hengstenberg R (1994) The halteres of the blowfly *Calliphora* II. Three-dimensional organization of compensatory reactions to real and simulated rotations. J Comp Physiol A 174: 695–708

Neumann TR, Bülthoff HH (2001) Insect inspired visual control of translatory flight. In: Proc EPSRC/BBSRC, International Workshop on Biologically Inspired Robotics 627–636.

O'Carroll D (1993) Feature-detecting neurons in dragonflies. Nature 362: 541–543

Pringle JWS (1948) The gyroscopic mechanism of the halteres of Diptera. Phil Trans R Soc Lond B 233: 347–384

Ramamurti R, Sandberg WC (2007) A computational investigation of the three-dimensional unsteady aerodynamics of *Drosophila* hovering and maneuvering. J Exp Biol 210: 881–896

Reichardt W, Poggio T (1975) A theory of the pattern induced flight orientation of the fly *Musca domestica* II. Biol Cybernetics 18: 69–80

Reichardt W, Poggio T (1976) Visual control of orientation behavior in the fly Part I. a quantitative analysis. Q Rev Biophys 9: 311–375

Reichardt W, Egelhaaf M, Guo A (1989) Processing of figure and background motion in the visual system of the fly. Biol Cybernetics 61: 327–345

Reiser MB, Dickinson MH (2003) A test bed for insect-inspired robotic control. Phil Trans R Soc Lond A 361: 2267–2285

Rosner R, Egelhaaf M, Warzecha A-K (2010) Behavioural state affects motion-sensitive neurones in the fly visual system. J Exp Biol 213: 331–338

Sane SP, Dieudonné A, Willis MA, Daniel TL (2007) Antennal mechanosensors mediate flight control in moths. Science 315: 863–866

Sherman A, Dickinson MH (2003) A comparison of visual and haltere-mediated equilibrium reflexes in the fruit fly *Drosophila melanogaster*. J Exp Biol 206: 295–302

Si A, Srinivasan MV, Zhang S (2003) Honeybee navigation: properties of the visually driven 'odometer'. J Exp Biol 206: 1265–1273

Srinivasan MV, Zhang SW, Altwein M, Tautz J (2000) Honeybee navigation: nature and calibration of the "odometer". Science 287: 851–853

Tammero LF, Dickinson MH (2002a) Collision-avoidance and landing responses are mediated by separate pathways in the fruit fly, *Drosophila melanogaster*. J Exp Biol 205: 2785–2798

Tammero LF, Dickinson MH (2002b) The influence of visual landscape on the free flight behavior of the fruit fly *Drosophila melanogaster*. J Exp Biol 205: 327–343

Tammero LF, Dickinson MH (2004) Spatial organization of visuomotor reflexes in *Drosophila*. J Exp Biol 207: 113–122

Tu MS, Dickinson MH (1996) The control of wing kinematics by two steering muscles of the blowfly, *Calliphora vicina*. J Comp Physiol A 178: 813–830

van Hateren JH, Schilstra C (1999) Blowfly flight and optic flight II. Head movements during flight. J Exp Biol 202: 1491–1500

Wagner H (1986) Flight performance and visual control of flight of the free-flying housefly (*Musca domestica* L.) II. Pursuit of targets. Phil Trans R Soc Lond B 312: 553–579

Wehner R (1981) Spatial vision in arthropods. In: Autrum H (ed) Handbook of sensory physiology. Springer, Berlin, Heidelberg, New York

Wolf R, Heisenberg M (1990) Visual control of straight flight in *Drosophila melanogaster*. J Comp Physiol A 167: 269–283

Zanker JM (1990) The wing beat of *Drosophila melanogaster* I. Kinematics. Phil Trans R Soc Lond B 327: 1–18

Zanker JM (1993) Theta motion: a paradoxical stimulus to explore higher order motion extraction. Vision Res 33: 553–569

Olfaction

III

Cuticular hydrocarbon sensillum for nestmate recognition in ants

10

Mamiko Ozaki, Midori Kidokoro-Kobayashi, Tetsutaro Hiraguchi

Contents

Abstract

Ants carry colony-specific blends of cuticular hydrocarbons (CHCs) on their exoskeleton as nestmate-discriminative pheromones and use them for social communication. To discriminate between such colony-specific blends of CHCs, Japanese carpenter ant workers, *Camponotus japonicus,* have a particular type of sensillum that responds to non-nestmate CHC mixtures but not to a nestmate's mixture. Once a worker ant detects non-nestmate blends of CHCs, its behavior changes drastically to aggression. Olfactory sensing of water-insoluble compounds like CHCs is mediated by lipophilic ligand carrier proteins. In the receptor lymph of insect chemosensilla those proteins have been discovered and categorized into the pheromone- and the general odorant-binding protein families, or the chemosensory protein family. These molecules are key in bridging chemical ecology and chemosensory physiology in insects, and are believed to be indispensable for the perireceptor event of not only olfactory but also gustatory systems concerned with feeding or mating. Their structure-function relationships have so far been studied in moths, fruit flies and other insects. *Cjap*CSP, the chemosensory protein found in *C. japonicus* CHC sensilla, is thought to introduce CHCs into the receptor lymph at the same mixing ratio at which they originally occur and to pass them on to the olfactory receptor neurons, which selectively respond to non-nestmate CHCs.

Mamiko Ozaki
Kobe University, Department of Biology
Rokkodai, Nada-ku
Kobe, Hyogo 657-8501, Japan
e-mail: mamiko@port.kobe-u.ac.jp

1. Introduction

Almost all animals including humans communicate with each other via various sensory systems. Human beings rely mainly on verbal communication. In the insect world, a single or a few pheromonal chemical components often do not convey all the necessary species-specific information, and so the chemical communication systems of insects often use mixtures of multiple compounds for con-specifics to interpret. In particular, social insects like honey bees or ants, which live in huge colonies, synthesize many kinds of pheromones and use them as "chemical words" for more complicated communication. Furthermore, the chemical information contained in pheromones, when sent to the central nervous system, is integrated with other information or referred to memorized information, until, finally, it reflects on behavior. In this article we focuse on the sensing of CHCs for nestmate recognition in ants. In various ant species, chemical analysis of the body surface materials suggests that the colony-specific blends of a multi-component CHC mixture act as the nestmate-discriminative pheromone. The CHC sensillum receives CHCs, colony specific mixtures of which are chemical signs shared by nestmates. Once a worker ant detects CHCs of non-nestmates via the CHC sensillum, its behavior becomes aggressive. The insect CHCs include straight chain, methyl, dimethyl, trimethyl branched and unsaturated components, and vary in volatility and fluidity, depending on the length of the carbon chain, unsaturated bonds and methyl branches. Past research perceived CHCs as simply a coating material preventing the small insect bodies from drying up. Subsequent research, however, has shown that in ants the CHCs and their mixtures are important cues for indicating belonging to nestmates and exclusiveness

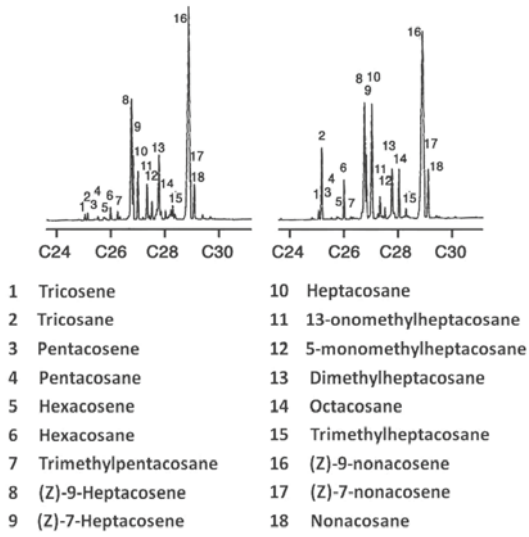


Fig. 1 Cuticular hydrocarbons (CHCs) of *Camponotus japonicus*. Gas chromatograms of CHCs of workers from two different colonies

for non-nestmates. In *C. japonicus*, the CHCs show 18 different components of 20–24 carbon chains, peak areas of which correspond to the relative amounts in a colony specific CHC blend (Fig. 1). Actually, in a multicomponent discriminant analysis, the mixing patterns of the CHCs of individuals of a particular colony belong to the same cluster whereas those of different colonies belong to different clusters. The pheromonal CHC components with long carbon chains used by ants are mostly non-volatile at the environmental temperatures prevailing for the ants, a fact which must be convenient for ants living in the closed space of the underground nest.

2. Chemosensory detection of cuticular hydrocarbons

2.1 Cuticular hydrocarbon perception in insects

For taste and olfactory sensing, the contact and the distant chemosensory sensilla are responsible for the reception of non-volatile and volatile chemicals, respectively. Thus, the non-volatile CHCs were thought to be received by taste sensilla (Ebbs and Amrein 2007). Recently, a male-specific CHC sex-pheromone inhibiting male-male courtship, was found to stimulate the bitter taste receptor neuron of the taste sensillum in *Drosophila melanogaster* (Lacaille et al. 2007). Prior to this finding, Ozaki et al. (2005) had discovered in the blowfly, *Phormia regina*, that a water-insoluble lipophilic material stimulated the bitter taste receptor neuron with the help of an odorant-binding protein distributed in the receptor lymph of taste sensilla. Moreover, Park et al. (2006) showed that another type of small protein, CheB42a, secreted in small subsets of taste sensilla on male forelegs, is required for a female-specific CHC to modulate male courtship behavior. Thus, the sex-pheromone CHCs consisting of a single species-specific or a few CHC components in *Drosophila* appear to stimulate taste sensilla.

Generally, the taste sensillum of ants possesses only a small number of receptor neurons for fundamental taste substances. Hence, the taste sensillum might not be suitable for the perception of CHC pheromones consisting of many components. For such multi-component pheromone reception, olfactory sensilla with many receptor neurons might be more suitable, even though CHC contact pheromones are not only water-insoluble but also non-volatile in most cases.

In the termite *Heterotermes tenuis* and the cockroach *Periplaneta americana*, it was suggested that some antennal sensory organs respond to CHCs as seen from the electroantennogram or the gas chromatography-electroantennographic detection (Batista-Pereira et al. 2003; Saïd et al. 2005). In the locust, *Schistocerca gregaria*, stimulation of antennae by CHCs induced synthesis of intracellular inositol tris-phosphate, which may act as a second messenger in some chemoreceptors (Heifetz et al. 1997). Honeybees can discriminate between different CHCs and learn to recognize and discriminate between them using their antennae (Getz and Smith 1987; Châline et al. 2005). However, the CHC sensitive sensilla have not yet been identified.

2.2 Electrophysiological recording from contact chemosensilla

In order to record action potentials from insect hair-shaped contact chemosensory sensilla, which typically have a single pore at their tip, Hodgson and Roeder (1956) originally developed the tip-recording method. When an indifferent electrode is inserted into the base of the target sensillum and a recording electrode containing electrolyte plus stimulant is in contact with the sensillar tip, the nervous responses are recorded simultaneously with the stimulation. Following the depolarization of the receptor membrane of the dendritic outer segment of a sensory neuron by chemical stimulation, a receptor current occurs at the dendritic outer segment (Ozaki and Tominaga1999). This induced inward current is largely driven by the transepithelial standing potential (generated by an electrogenic K^+ pump of auxiliary cells facing the receptor lymph space; see also Thurm and Küppers 1980) and reaches the impulse generating site. The impulses are generated by the bulk of outward current, which would occur around the ciliary segment between the outer and the inner segments of the receptor neuron (Ozaki and

Tominaga 1999). The receptor membrane depolarization is the earliest recorded electrical event, followed by impulse generation. As long as the impulse frequency is linearly related to the receptor potential (Ozaki and Amakawa 1990), it is used as a quantitative indicator of neuronal activity of the stimulated receptor neuron.

2.3 Functional identification of the ant cuticular hydrocarbon sensillum

Because of the antennation behavior shown by the ants when inspecting other ants, the CHC-sensitive sensilla were expected to be found on the antennae. However, CHCs are too lipophilic to be dissolved in an aqueous electrolyte solution. This is one of the biggest problems when preparing the stimulus in search of the CHC sensillum on the antenna using electrophysiology. In order to solve this problem, we first tried to dissolve the CHCs in the electrolyte by adding a mild detergent, and then alternatively dissolved them with a chemosensory protein. The *Cjap*CSP (chemosensory protein) was discovered in the CHC sensillum of *C. japonicus* as a lipophilic carrier protein (Ozaki et al. 2005). It was later proposed that CHCs be transported to the receptor membranes across the aqueous receptor lymph in the CHC sensillum. When a recording electrode filled with non-nestmate CHC solution contacted the target CHC sensillum, a vigorous discharge of impulses with a variety of amplitudes was observed indicating the presence of several sensory cells.

The hair shaft of the antennal contact chemosensory sensillum sensitive to CHCs was 15 µm long and 4 µm in diameter, and morphologically identified as an olfactory sensillum (arrow in Fig. 2A). It shows numerous tiny olfactory pores (Fig. 2B), through which the stimulus CHCs are believed to penetrate inside the sensillum to subsequently stimulate the receptor neurons (Fig. 2C). As a result, the tip-recording method, which was originally developed for taste response recording, was also applicable to the CHC sensillum without a gustatory tip pore.

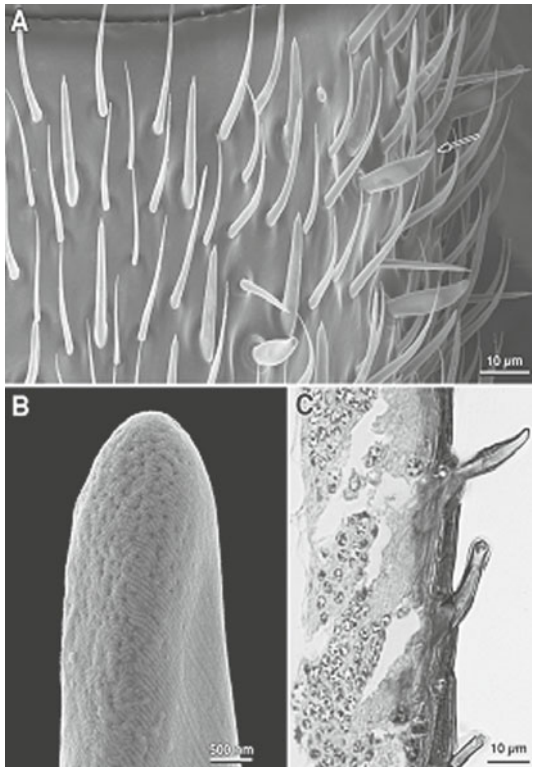


Fig. 2 CHC sensillum of *C. japonicus* (Photos by M. Iwasaki). **A** Scanning electron micrograph of antennal surface. **B** High magnification of tip of a CHC sensillum. **C** Transmission electron micrograph including a CHC sensillum (top)

Transmission electron microscopic analysis revealed the presence of about 130 receptor neurons in a single CHC sensillum (Ozaki, unpublished data). Vosshall (2001) and Couto et al. (2005) proposed that a single receptor neuron responds to multiple odorants and that one odorant stimulates several receptor neurons. Following this proposal even a single CHC component might induce impulse generation in several receptor neurons within the same CHC sensillum. Thus, when a CHC sensillum of an ant is stimu-

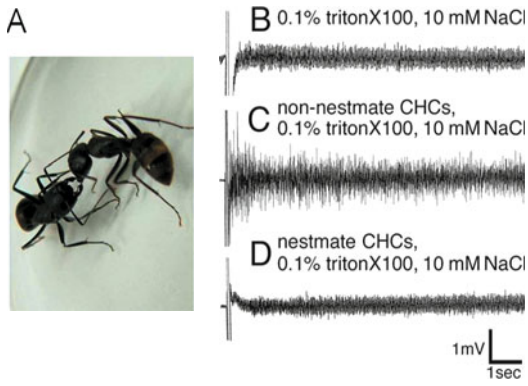


Fig. 3 Electrophysiological response of CHC sensillum. **A** Aggressive behavior of *C. japonicus*. **B** Responses to solvent, **C** response to non-nestmate CHCs and **D**, to nestmate CHCs

lated with a CHC mixture extracted from the body surface of ants, several different receptor neurons should generate impulses. As expected, in experiments with various combinations of colonies, to which the test ant and the CHC donor ant belonged, sometimes (but not always) vigorous impulse discharge of the stimulated CHC sensillum was observed (Fig. 3). Surprisingly, the CHC sensillum did discriminate between the non-nestmate CHC mixture and the nestmate CHC mixture. Indeed, the CHC sensillum selectively responded to the non-nestmate CHC mixtures (Fig. 3 C) with action potentials that triggered aggressive behavior (Fig. 3 A).

3. Perireceptor events in the cuticular hydrocarbon sensillum

3.1 Lipophilic ligand-carrier proteins involved in the perireceptor events

In insect chemosensilla, two classes of the carrier proteins, odorant-binding proteins and chemosensory proteins are found. Those small amphipathic proteins are secreted from auxiliary cells into the receptor lymph surrounding the receptor cell membranes. Even lipophilic pheromones and odorants have to reach the receptor cell membranes across the aqueous receptor lymph. Hence, odorant-binding proteins or chemosensory proteins were predicted to be carrying and transporting lipophilic ligands to the receptor membranes (Vogt and Riddiford 1981; Vogt et al. 1990; Pelosi 1998; Steinbrecht 1998; Jacquin-Joly et al. 2001; Calvello et al. 2003; Ozaki et al. 2005; Kaissling 2009) Odorant-binding proteins or chemosensory proteins share some characteristics: small size (11–18 kDa), high water solubility and reversible binding activity for small lipophilic molecules. There is, however, a low similarity among their amino acid sequences (Pelosi et al. 2006). Odorant-binding proteins and chemosensory proteins commonly contain α-helical domains, but are folded differently.

The odorant-binding proteins are widely differentiated and distributed across and within species, with identical amino acid residues around 10–15 % (Pelosi et al. 2006). Their signature is represented by a pattern of six cysteins in conserved positions that in the native protein form three interlocking disulphide bridges. These disulfide bridges give the molecule a rigid structure so that the helices cannot move freely. Instead, only the N- and C-terminal parts of the protein are flexible and can move to open and close the ligand-binding cavity (Prestwich et al. 1995; Leal et al. 1999; Scaloni et al. 1999).

Odorant-binding proteins are further classified into pheromone-binding proteins and general odorant-binding proteins (Laue et al. 1994; Steinbrecht et al. 1995; Krieger et al. 1996; Zhang et al. 2001).

The pheromone-binding proteins in the male moth antennae have considerable binding selectivity for the major constituent of the female sex pheromones in each species (Leal et al. 2005). In the three-dimen-

sional folding of pheromone-binding protein of *Bombyx mori*, the position of its ligand, bombykol, was identified inside the binding cavity, which is located at the core of the protein (Shandler et al. 2000). The lipophilic ligand in the binding pocket can be perfectly separated from the hydrophilic receptor lymph (Shandler et al. 2000). The C-terminal region of the *Bombyx mori* pheromone-binding protein folds into an α-helical segment and fits into the ligand-binding cavity at low pH, although it has no definite structure at neutral pH (Wojtasek and Leal 1999). This pH-dependent structural change of the C-terminal was thought to occur in the local acidic environment close to the dendritic membrane. It was then hypothesized that the α-helical C-terminal of *Bombyx mori* pheromone-binding protein could push the pheromone molecule out of the ligand-binding cavity and pass on the pheromone to the receptor protein on the dendritic membrane (Leal 2005).

The general odorant-binding proteins, which have mainly been found in the female moth antennae, seem to carry odorants of plant volatiles (Steinbrecht et al. 1995). In the honeybee, *Apis mellifera*, ASP2 (AmelOBP2) expressed in the antennae of workers and drones can bind odorant molecules of floral scents such as 1,8-cineol, 2-isobutyl-3-methoxypyrazine, 2-heptanone and isoamylacetate but not pheromone components (Danty et al. 1997, 1999; Briand et al. 2001).

On the other hand, the chemosensory proteins are conserved with 40–50 % amino acid identity even between phylogenetically distant species (Pelosi et al. 2006). Their signature consists of four cysteins connected by disulphide bonds between adjacent cystein residues, resulting in the formation of two small loops of eight and four amino acids. Chemosensory proteins may not have so rigid a structure as odorant-binding proteins and the ligand-binding cavities of chemosensory proteins may be more flexible than those of odorant-binding proteins (Picimbon 2003). The ligand-binding pocket of chemosensory proteins is imagined to be a sort of hydrophobic tunnel, which can accommodate a long carbon chain molecule (Lartigue et al. 2002, Mosbah et al. 2003; Campanacci et al. 2003).

Originally, odorant-binding proteins were discovered in lepidopteran antennae and considered to be involved in perireceptor events around odorant receptor membranes, whereas chemosensory proteins (yet undefined in their function) were found not only on antennae but also on other parts of the exoskeleton where chemosensory sensilla are located (Steinbrecht 1998; Shanbhag et al. 2001). However, antennal specificity of odorant-binding proteins has not been verified, since an odorant-binding protein was found in the labellar taste sensilla of the blowfly, *Phormia regina* (Ozaki et al. 1995). In some social hymenopterans, odorant-binding protein is expressed in body parts other than the antennae, such as legs and wings (Calvello et al. 2003, 2005). Among the Formicidae, the Argentine ant, the Japanese carpenter ant and the red fire ant are known to have antenna specific chemosensory proteins, which may carry lipophilic CHC molecules to the receptor membranes across the receptor lymph (Fig. 4) (Ishida et al. 2002; Ozaki et al. 2005; Leal and Ishida 2008).

3.2 Ligand binding of lipophilic ligand carrier proteins

Ligand-binding studies of odorant-binding proteins and chemosensory proteins have been carried out in different insect species which have led to different results (Pelosi et al. 2006). In the social wasp, *Polistes dominulus*, both odorant-binding protein and chemosensory protein have good affinity for differently classified long chain compounds,

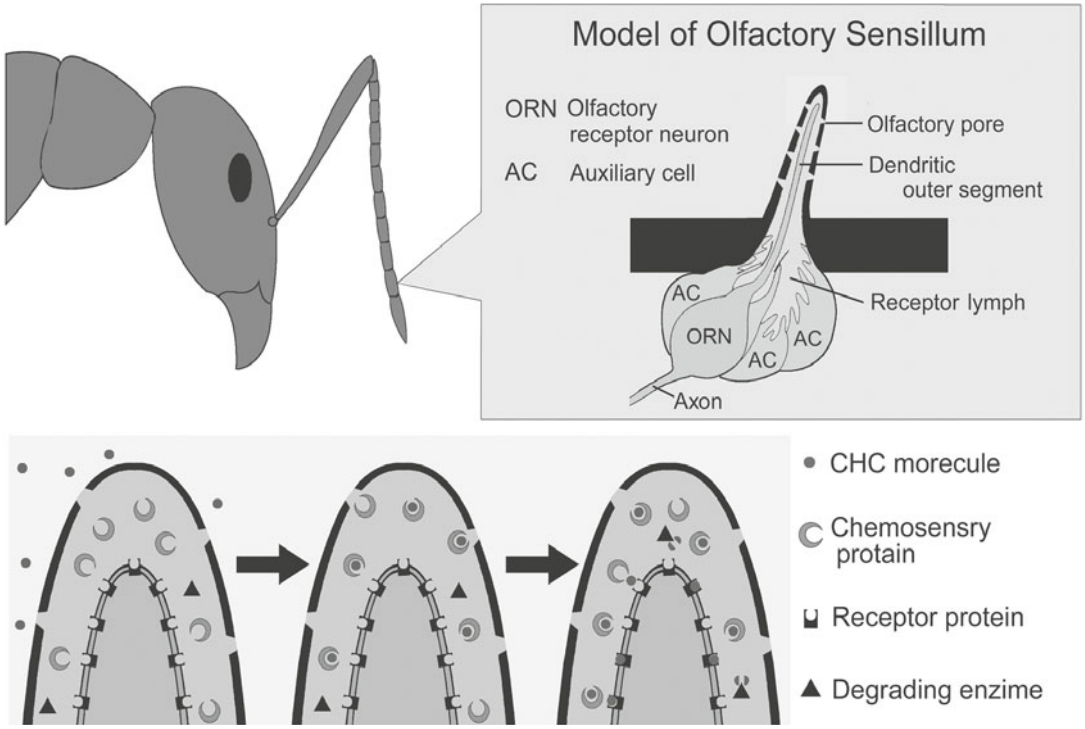


Fig. 4 Schematic of the CHC sensillum. CHC molecule penetrating through the sidewall pore is carried by chemosensory protein (CSP) to the receptor membrane across the sensillar lymph

such as acids, esters and amines (Calvello et al. 2003). The odorant-binding protein is commonly found in sensilla of the antennae, head, wings and legs of all castes and ages. The chemosensory protein is mainly expressed in the antennal sensilla of workers and males, but not in those of females (except for workers). It is also expressed in sensilla on the legs (Calvello et al. 2003, 2005). Chemosensory protein exhibited a broader binding spectrum for linear saturated amides, alcohols, and carboxylic acids with 12–18 carbons than odorant-binding protein did for saturated amides with 12–16 carbons and saturated straight chain alcohols of 12–18 carbons. In the honeybee, *Apis mellifera*, ASP3c was found in sensilla of the antennae, wings and legs and is proposed to be a general lipid carrier protein. It binds several fatty acids and fatty acid methyl esters: myristic acid, palmitic acid, stearic acid, C16-methyl ester and C18-methyl ester, the latter two of which are components in the brood pheromone blend (Briand et al. 2002). In the locust, *Schistocerca gregaria*, chemosensory proteins are found in the taste sensilla rather than in the olfactory sensilla (Angeli et al. 1999). CSP-sg4 binds various plant volatiles, carboxylic acids and linear alcohols with 12, 14 and 18 carbon atoms, but does not bind aggregation pheromones (Ban et al. 2002). In *Mamestra brassicae*, CSPMbraA6 in the antennae and pheromone gland binds volatile pheromone components and brominated alkyl alcohols or fatty acids with 12–18 carbons (Jacquin-Joly et al. 2001; Lartigue et al. 2002).

At the present time, only CSPMbraA6 of *Mamestra brassicae* is structurally characterized (Lartigue et al. 2002, Campanacci et al. 2003; Mosbah et al. 2003). The molecular structure was analyzed in the binding of

a pheromone component, Z11–16: Ac, and showed large conformational changes after ligand-binding. The binding pocket is reported to bind up to three 12-bromo-dodecanol molecules (Lartigue et al. 2002). It is as yet unclear whether the chemosensory protein binds ligands with positive cooperative kinetics, but precise investigation of the three dimensional structure of chemosensory proteins in various insect species would suggest more about the function of this lipophilic ligand carrier protein family.

Recently, in the fruit fly *Drosophila melanogaster*, CheB42a was found in the taste sensilla on the front legs of the males (Xu et al. 2002; Starostina et al. 2008). CheB42a is secreted into the inner lumen of the particular pheromone-sensing taste hairs (Starostina et al. 2008). The CheB proteins, including CheB42a, are similar to MD-like proteins (MD, myeloid differentiation protein). MD-like proteins, which are found in all eukaryotes, form the ML family. The ML family is a lipid-binding protein family different from the odorant-binding protein or the chemosensory protein family. These proteins are soluble proteins of 150–200 amino acids with two or more disulfide bonds and have signal peptides at their N-terminal. Although not much has been reported about the molecular mechanisms of MD-like proteins for hydrocarbon pheromone detection, it is suggested that the protein is concerned with the perception of hydrocarbons used as contact sex pheromones (Starostina et al. 2009). Indeed, in *Drosophila melanogaster*, female specific CHC pheromones play essential roles in triggering and modulating mating behavior.

4. Chemoreception of cuticular hydrocarbons and related behavior expression

4.1 Role of chemosensory proteins in the ant CHC sensillum

For nestmate recognition in the Japanese carpenter ant, *Camponotus japonicus*, the worker ants distinguish the colony specific CHC profiles consisting of over 18 compounds using their antennal chemosensory sensilla sensitive to CHCs (Ozaki et al. 2005). In order to prove that *Cjap*CSP, the chemosensory protein of *C. japonicus*, functions as a CHC carrier protein in this particular type of antennal chemosensory sensilla, CHCs extracted from the ant body surface were adsorbed onto the glass wall of a test tube by evaporating organic solvents. Subsequently, the buffer solution containing *Cjap*CSP was added to the tube, so that the CHCs in the *Cjap*CSP-binding form were dissolved into the buffer. Then, the CHCs were extracted with a small amount of organic solvent and quantitatively analyzed by gas-chromatography. When an aqueous buffer containing *Cjap*CSP was used instead of intrinsic receptor lymph, *Cjap*CSP nonspecifically bound every CHC component, keeping the same CHC mixture pattern as that found on the ant body surface, regardless of the colony specificity of the CHC pattern (Fig. 5). The buffer containing bovine serum albumin instead of *Cjap*CSP dissolved very little CHCs and did not reproduce the original CHC pattern.

By using a *Cjap*CSP-containing electrolyte solution, electrophysiological stimulus solutions using CHC mixtures of various colonies were prepared, and the impulses in response to those CHC mixtures were recorded from the CHC-sensilla. Nestmate CHCs never stimulated the CHC-sensilla but

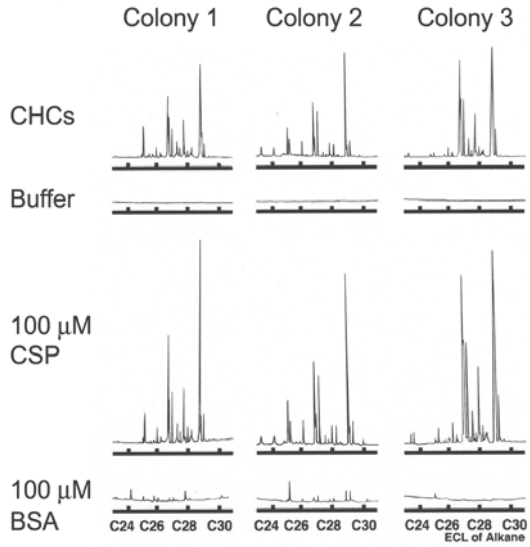


Fig. 5 Dissolving of CHCs in aqueous solutions. Original CHC patterns from three different colonies (top) were dissolved in phosphate buffer, 100 µM chemosensory protein (CSP) in phosphate buffer or 100 µM bovine serum albumin (BSA) in phosphate buffer. CHCs were not detected in phosphate buffer. A small amount of CHCs were detected in phosphate buffer containing BSA but not at the original ratio

non-nestmate CHCs did as mentioned above (Fig. 3). This implies that the nestmate and non-nestmate discrimination primarily occurs at the sensillar level already. Thus, the strict nestmate-non-nestmate discrimination based on the colony specific CHC patterns is not possible without *Cjap*CSP.

4.2 Chemical communication by CHCs in ant society

We mentioned that the body surface chemicals of insects contain a large proportion of CHCs, which function as important components in chemical communication (Howard 1993). Ant species, especially, rely heavily on CHCs to chemically communicate the membership of colonies, caste, sex and physiological status. The CHC components differ among species, and their mixture profiles are characteristically different among colonies, but rather resemble each other in the individual worker ants in each colony. Colony members usually share similar CHC profiles, used as chemical sign for nestmate recognition (Bonavita-Cougourdan et al. 1987; Yamaoka 1990; Howard 1993; Vander Meer and Morel 1998; Howard and Blomquist 2005). Thus, based on the chemical signatures of CHC mixtures, nestmate recognition underlies various forms of social behavior in ants. The hypothesis claiming the essential role of CHCs in the social behaviors of ants, has been supported by experimental evidences derived from a variety of behavioral bioassays or correlation studies in which cuticular compounds were removed and reapplied via solvent extraction (Bonavita-Cougourdan et al. 1987; Howard 1993; Breed 1998; Vander Meer and Morel 1998; Guerreri and d'Ettore 2008). This hypothesis was further supported by a comparison of the pheromone activity between natural and artificial hydrocarbon blends using synthetic hydrocarbon components (Lahav et al. 1999; Akino et al. 2002, 2004; Ozaki et al. 2005; Martin et al. 2008a).

Nestmate recognition can be evaluated by assessing the aggressiveness of workers against non-nestmate workers. Such colony discriminative aggression was observed not only against living non-nestmate workers but also against glass bead dummies, onto which non-nestmate CHCs were deposited. By using glass beads as surrogate ants, the stimulation to trigger aggressive behaviors was well assessed not only qualitatively, but also quantitatively. Aggressive behavior toward non-nestmates is regulated by the CHC profiles consisting of over 18 compounds. The ants can distinguish the CHC profiles of non-nestmates from that of nestmates by a single sweep of their antennae (Wilson 1971; Tanner 2008).

According to the CHC-dose-dependent

behavioral experiments in *C. japonicus*, we found that the aggression threshold of individuals broadly varies over 3-log units from 10^{-4} to 10^{-1} ant equivalent of non-nestmate CHCs. However, the median thresholds of different ant populations, even from different colonies, are all around 0.005 ant equivalent. Thus, considering the broad variation of the aggression threshold of individuals, these might have various CHC sensitivities at either the sensillar level or at higher neuronal levels. It should be explained at which level and how the expression of aggressive behavior toward non-nestmates is regulated. When a single sensillum is stimulated with a non-nestmate CHC blend of multiple CHC components, multiple receptor neurons respond with impulses of different shapes originating in many sensory cells. By measuring the impulse frequency, the sensillar sensitivity to CHCs can be evaluated. When the antenna sweeps the body surface of a non-nestmate, many sensilla are stimulated. The impulses from the stimulated sensilla are summed in a primary olfactory center, the antennal lobe (Nishikawa et al. 2008) and then the aggression-inducing neuronal signal is sent to the projection neurons (Kirschner et al. 2006).

Conclusions

In animal societies, chemical communication plays an important role in conflict and cooperation. In ants colony-specific CHC complexes serve as chemical cues for the discrimination between nestmates and non-nestmates. The discovery of an antennal sensillum, which responds to the CHCs of non-nestmates but not to those of nestmates, shows that the mechanisms involved in the build-up of social aggression against non-nestmates start way out in the periphery. The molecular mechanism underlying the perireceptor and receptor events in the CHC sensillum considerably influence the behavioral response.

One of the remaining problems to be solved by future research is the mechanism of the remarkable filter function in the CHC sensillum, which blocks the transfer of chemical information from nestmates but passes on that from non-nestmates. Moreover, it is still puzzling that the CHC sensillum of *C. japonicus* potentially may distinguish between 2^{130} different patterns by an on/off combination of its 130 receptor neurons. Obviously, this number of receptor neurons in a sensillum is not needed to recognize CHC blends consisting of 18 components only, so possibly the sensillum needs to respond to CHCs not only of conspecifics but also of other species.

List of abbreviations

CHC cuticular hydrocarbon; *Cjap Camponotus japonicus*; CSP chemosensory protein

References

Akino T, Terayama M, Wakamura S, Yamaoka R (2002) Intraspecific variation of cuticular hydrocarbon composition in *Formica japonica* Motschoulsky (Hymenoptera: Formicidae). Zool Sci 19: 1155–1165

Akino T, Yamamura K, Wakamura S, Yamaoka R (2004) Direct behavioral evidence for hydrocarbons as nestmate recognition cues in *Formica japonica* (Hymenoptera: Formicidae). Appl Entomol Zool 39: 381–387

Angeli S, Ceron F, Scaloni A, Monti M, Monteforti G, Minnocci A, Petacchi R, Pelosi P (1999) Purification, structural characterization, cloning and immunocytochemical localization of chemoreception proteins from *Schistocerca gregaria*. Eur J Biochem 262: 745–754

Ban L, Zhang L, Yan Y, Pelosi P (2002) Binding properties of a locust's chemosensory protein. Biochem Biophys Res Comm 293: 50–54

Batista-Pereira LG, dos Santos MG, Corrêa A G, Fernandes J B, Arab A, Costa-Leonardo A- M, Dietrich C R R C, Pereira D A, Bueno O C (2003) Cuticular hydrocarbons of *Heterotermes tenuis* (Isoptera: Rhinotermitidae): Analyses and electrophysiological studies. Z Naturforsch 59c: 135–139

Bonavita-Cougourdan A, Clément J-L, Lange C (1987) Nestmate recognition: The role of cuticular hydrocarbons in the ant *Camponotus vagus.* Scop J Entomol Sci 22: 1–10

Breed M D (1998) Chemical cues in kin recognition: criteria for identification, experimental approaches, and the honey bee as an example. In: Van der Meer R K, Breed M D, Espelie K E, Winston M L (eds.) Pheromone communication in social insects: ants, wasps, bees and termites. Westview Press, Boulder, pp 57–78

Briand L, Nespoulous C, Huet J -C, Takahashi M, Pernollet, J -C (2001) Ligand binding and physico-chemical properties of ASP2, a recombinant odorant-binding protein from honeybee (*Apis mellifera* L.). Eur J Biochem 268: 752–760

Briand L, Nespoulous C, Huet J -C, Takahashi M, Pernollet J -C (2002) Characterization of a chemosensory protein (ASP3c) from honeybee (*Apis mellifera* L.) as a brood pheromone carrier. Eur J Biochem 269: 4586–4596

Calvello M, Brandazza A, Navarrini A, Dani F R, Turillazzi S, Felicioli A, Pelosi P (2005) Expression of odorant-binding proteins and chemosensory proteins in some Hymenoptera. Insect Biochem Molec Biol 35: 297–307

Calvello M, Guerra N, Brandazza, A, D'Ambrosio C, Scaloni A, Dani F R, Turillazzi S, Pelosi P (2003) Soluble proteins of chemical communication in the social wasp *Polistes dominulus*. Cell Mol Life Sci 60: 1933–1943

Campanacci V, Lartigue,A, Hallgerg M M, Jones T A, Giudici-Orticoni M-T, Tegoni M, Cambillau C (2003) Moth chemosensory protein exhibits drastic conformational changes and cooperativity on ligand binding. Proc Natl Acad Sci 100: 5069–5074

Châline N, Sandoz J-C, Martin S J, Ratnieks F L W, Jones G R (2005) Learning and discrimination of individual cuticular hydrocarbons by honeybees (*Apis mellifera*). Chem Senses 30: 327–335

Couto A, Alenius M, Dickson B J (2005) Molecular, anatomical, and functional organization of the *Drosophila* olfactory system. Curr Biol 15**:** 1535–1547

Danty E, Briand L, Michard-Vanhée C, Perez V, Arnold G, Gaudemer O, Huet D, Huet, J-C, Ouali C, Masson C, Pernollet J-C (1999) Cloning and expression of a queen pheromone-binding protein in the honeybee: an olfactory-specific, developmentally regulated protein. J Neurosci 19**:** 7468–7475

Danty, E, Michard-Vanhée C, Huet J-C, Genecque E, Pernollet J-C, Masson C (1997) Biochemical characterization, molecular cloning and localization of a putative odorant binding protein in the honeybee *Apis mellifera* L. FEBS let 414: 595–598

Ebbs M L, Amrein H (2007) Pheromone and taste perception in the fruit fly, *Drosophila melanogaster.* Pflügers Arch-Eur J Physiol 454: 735–747

Getz W M, Smith K B (1987) Olfactory sensitivity and discrimination of mixtures in the honeybee *Apis mellifera.* J Comp Physiol A 160: 239–245

Guerreri F J, d'Ettore P (2008) The mandible opening response: quantifying aggression elicited by chemical cues in ants. J Exp Biol 211: 1109–1113

Heifetz Y, Boekhoff I, Breer H, Appplebaum S W (1997) Cuticular hydrocarbons control behavioural phase transition in *Schistocerca gregaria* nymphs and elicit biochemical responses in antennae. Insect Biochem Molec Biol 27: 563–568

Hodgson E S, Roeder K D (1956) Electrophysiological studies of arthropod chemoreception. I. General properties of the labellar chemoreceptors of Diptera. J Cell Comp Physiol 48: 41–74

Howard R W (1993) Cuticular hydrocarbons and chemical communication. In: Stanley-Samuelson D W, Nelson D R (eds.) Insect lipids: chemistry, biochemistry and biology. University of Nebraska Press, Lincoln and London, pp 179–226

Howard R W, Blomquist G J (2005) Ecological, behavioral, and biochemical aspects of insect hydrocarbons. Annu Rev Entomol 50: 371–393

Ishida Y, Chiang V, Leal W S (2002) Protein that makes senses in the Argentine ant. Naturwissenschaften 89: 505–507

Jacquin-Joly E, Vogt R G, Francois M-C, Nagnan-Le Meillour P (2001) Functional and expression pattern analysis of chemosensory proteins expressed in antennae and pheromonal gland of *Mamestra brassicae*. Chem Senses 26: 833–844

Kaissling K-E (2009) Olfactory perireceptor and receptor events in moths: a kinetic model revised. J Comp Physiol A 195: 895–922

Kirschner S, Kleineidam C J, Zube C, Rybal J, Grunewald B, Rossler W (2006) Dual olfactory pathway in the honeybee, *Apis mellifera*. J Comp Neurol 499: 933–952

Krieger J, von Nickisch-Rosenegk E, Mameli M, Pelosi P, Breer H (1996) Binding proteins from the antennae of *Bombyx mori.* Insect Biochem Molec Biol 3: 297–307

Lacaille F, Hiroi M, Twele R, Inoshita T, Umemoto D, Manière G, Marion-Poll F, Ozaki M, Francke W, Cobb M, Everaerts C, Tanimura T, Ferveur J-F (2007) An inhibitory sex pheromone tastes bitter for *Drosophila* males. PLoS ONE 2: e661

Lahav S, Soroker V, Hefetz A, Van der Meer R K (1999) Direct behavioral evidence for hydrocarbons as ant recognition discriminators. Naturwissenschaften 86: 246–249

Lartigue A, Campanacci V, Roussel A, Larsson A M, Jones T A, Tegoni M, Cambillau C (2002) X-ray structure and ligand binding study of a moth chemosensory protein. J Biol Chem 277: 32 094–32 098

Laue M, Steinbrecht R A, Ziegelberger G (1994) Immunocytochemical localization of general odorant binding protein in olfactory sensilla of the silkworm *Antheraea polyphemus*. Naturwissenschaften 81: 178–180

Leal W S (2005) Pheromone reception. Top Curr Chem 240: 1–36

Leal W S, Chen A M, Erickson M L (2005) Selective and pH-dependent binding of a moth pheromone to a pheromone-binding protein. J Chem Ecol 31: 2493–2499

Leal W S, Ishida Y (2008) GP-9s are ubiquitous proteins unlikely involved in olfactory mediation of social organization in the red imported fire ant, *Solenopsis invicta*. PLoS ONE 3: e3762

Leal W S, Nikonova L, Peng G (1999) Disulfide structure of the pheromone binding protein from the silkworm moth, *Bombyx mori*. FEBS let 464: 85–90

Martin S J, Helanterä H, Drijfhout F P (2008a) Colony-specific hydrocarbons identify nest mates in two species of *Formica* ant. J Chem Ecol 34: 1072–1080

Martin S J, Vitikainen E, Helanterä H, Drijfhout F P (2008b) Chemical basis of nest-mate discrimination in the ant *Formica exsecta*. Proc R Soc Lond B 270: 153–158

Mosbah A, Campanacci V, Lartigue A, Tegoni M, Cambillau C, Darbon H (2003) Solution structure of a chemosensory protein from the moth *Mamestra brassicae*. Biochem J 369: 39–44

Nishikawa M, Nishino H, Misaka Y, Kubita M, Tsuji E, Satoji Y, Ozaki M, Yokohari F (2008) Sexual dimorphism in the antennal lobe structure of the ant, *Camponotus japonicus*. Zool Sci 25: 195–204

Ozaki M, Amakawa T (1990) Adaptation-promoting effects of IP_3, Ca^{2+} and phorbol ester on the sugar taste receptor cell of the blowfly, *Phormia regina*. J Gen Physiol 100: 867–879

Ozaki M, Morisaki K, Idei W, Ozaki K, Tokunaga F (1995) A putative lipophilic stimulant carrier protein commonly found in the taste and olfactory system: A unique member of the pheromone-binding protein superfamily. Eur J Biochem 230: 298–308

Ozaki M, Takahara T, Kawahara Y, Wada-Katsumata A, Seno K, Amakawa T, Yamaoka R, Nakamura T (2003) Perception of noxious compounds by contact chemoreceptors of the blowfly, *Phormia regina*: Putative role of an odorant-binding protein. Chem Senses 28: 349–359

Ozaki M, Tominaga Y (1999) Contact Chemoreceptors. In: Eguchi E, Tominaga Y (eds.) Atlas of arthropod sensory receptors. Springer-Verlag, Tokyo, pp 143–154

Ozaki M, Wada-Katsumata A, Fujikawa K, Iwasaki M, Yokohari F, Satoji Y, Nishikawa T, Yamaoka R (2005) Ant nestmate/non-nestmate discrimination by a chemosensory sensillum. Science 309: 311–314

Park S K, Mann K J, Lin H, Starostina E, Kolski-Andreaco A, Pikielny C W (2006) A *Drosophila* protein specific to pheromone-sensing gustatory hairs delays males' copulation attempts. Curr Biol 16: 1154–1159

Pelosi P (1998) Odorant-binding proteins: Structural aspects. Ann N Y Acad Sci 855: 281–293

Pelosi P, Zhou J-J, Ban L P, Calvello M (2006) Soluble proteins in insect chemical communication. Cell Mo Life Sci 63: 1658–1676

Picimbon J-F (2003) Biochemistry and evolution of OSD and OBP proteins. In: Blomquist G J, Vogt R G (eds) Pheromone biochemistry and molecular biology. New ed. Academic Press, New-York, pp 539–566

Prestwich G D, Du G, LaForest S (1995) How is pheromone specificity encoded in proteins? Chem Senses 20: 461–469

Saïd I, Gaertner C, Renou M, Rivault C (2005) Perception of cuticular hydrocarbons by the olfactory organs in *Periplaneta americana* (L.) (Insecta, Disctyoptera). J Insect Physiol 51: 1384–1389

Scaloni A, Monti M, Angeli S, Pelosi P (1999) Structural analysis and disulfide-bridge pairing of two odorant-binding proteins from *Bombyx mori*. Biochem Biophys Res Comm 266: 386–391

Shanbhag S R, Hekmat-Scafe Kim M S, Park S K, Carlson J R, Pikielny C, Smith D P, Seinbrecht R A (2001) Expression mosaic of odorant-binding proteins in *Drosophila* olfactory organs. Microsci Res Tech 55: 297–306

Shandler B H, Nikonova L, Leal W S, Clardy J (2000) Sexual attraction in the silkworm moth: struc-

ture of the pheromone-binding-protein-bombykol complex. Chem Biol 7: 1434–151

Starostina E, Xu A, Lin H, Pikiely C W (2009) A *Drosophila* protein family implicated in pheromone perception is related to Tay-Sachs GM2-activator protein. J Biol Chem 284**:** 585−594

Steinbrecht R A (1998) Odorant binding proteins: expression and function. Ann N Y Acad Sci 855: 323−332

Steinbrecht R A, Laue M, Ziegelberger G (1995) Immunolocalization of pheromone binding protein and general odorant-binding protein in olfactory sensilla of the silk moths *Antheraea* and *Bombyx*. Cell Tissue Res 282: 203−217

Tanner C J (2008) Aggressive group behaviour in the ant *Formica xerophila* is coordinated by direct nestmate contact. Anim Behav 76: 1335–1341

Thurm U, Küppers J (1980) Epithelial physiology of insect sensilla. In: Locke M, Smith D (eds.) Insect biology in the future. Academic Press, New York, pp 735–763

Van der Meer R K, Morel L (1998) Nestmate recognition in ants. In: Van der Meer R K, Breed M D, Espelie K E, Winston M L (eds.) Pheromone communication in social insects: ants, wasps, bees and termites. Westview Press, Boulder, pp 79–103

Vosshall L B (2001) The molecular logic of olfaction in *Drosophila*. Chem Senses 26: 207−213

Vogt R G, Riddiford L M (1981) Pheromone binding and interaction by moth antenna. Nature 293: 161–163

Vogt R G, Rybczynshi R, Lerner M R (1990) The biochemistry of odorant reception and transduction. In: Schild D (ed.) Chemosensory information processing. Springer-Verlag, Berlin, pp 33–76

Wilson E O (1971) The insect societies. Belknap Press of Harvard University Press, Cambridge, Massachusetts

Wojtasek H, Leal W S (1999) Conformational change in pheromone-binding protein from *Bombyx mori* induced by pH and by interaction with membranes. J Biol Chem 274: 30950−30956

Xu A, Park S K, D'Mello S, Kim E, Wang and Pikielny C (2002) Novel genes expressed in subsets of chemosensory sensilla on the front legs of male *Drosophila melanogaster.* Cell Tissue Res 307: 381–192

Yamaoka R (1990) Chemical approach to understanding interaction among organisms. Physiol Ecol Japan 27: 31–52

Zhang S, Maida R, Steinbrecht R A (2001) Immunolocalization of odorant-binding-proteins in noctuid moths (Insecta, Lepidoptera). Chem Senses 26: 885−896

Fluid mechanical problems in crustacean active chemoreception

11

DeForest Mellon Jr., Matthew A. Reidenbach

Contents

Abstract

Aquatic crustaceans detect odorants in their fluid environment using batteries of microscopic cuticular sensilla, termed aesthetascs, which are arrayed along their antennules. Because these structures are tiny they operate at small Reynolds numbers, implying that fluid flow around and within the sensor arrays is laminar. Access of odorants to the surface of the aesthetascs therefore occurs primarily via molecular diffusion, a process that in most crustacean species appears to be enhanced by antennular flicking behavior. Depending upon the density of the sensor array and the size and structure of the individual aesthetascs, flicking may enhance 'leakiness' of the flow through the sensor array, thereby decreasing the distance over which odorant-laden fluid must be molecularly diffused to the surface of the individual aesthetascs. Here we review theoretical considerations of the fluid mechanics involved with odorant access to olfactory sensors on the antennules of aquatic crustaceans, including the nature of some remaining unanswered questions, and a brief comparison with the situation in crustaceans that exhibit a terrestrial lifestyle.

DeForest Mellon Jr.
University of Virginia, Department of Biology
Gilmer Hall, McCormick Road
Charlottesville, VA 22901, USA
e-mail: dm6d@cms.mail.virginia.edu

1. Introduction

Crustaceans have a central nervous system largely devoted to capturing and analyzing the olfactory environment. Much of their brain (as much as 40 % in some species, but 50 % or more in terrestrial hermit crabs) is devoted to the analysis of incoming olfactory signals, integrating those signals with inputs from visual, mechanical and taste receptors on the head appendages, and to generating behaviors that enhance their efficiency in detecting olfactory signals. There are significant gaps in our understanding of these processes at the cellular and subcellular level. For example, crustaceans have

contact chemoreceptors on their legs as well as on their antennae. A major aspect of active chemoreception in crustaceans involves probing the substrate with their dactyls, probably integrating the chemosensory input with the simultaneous input from contact mechanoreceptor neurons housed within the same sensilla. We have little understanding of how this integration between this particular subset of chemical sensors and mechanoreceptors occurs and how it is used to aid in detecting and processing odorant signals. We also have little understanding of how information obtained from contact chemoreceptor sensilla located on their dactyls integrates with that from the olfactory sensilla located on their antennules. A separate but related problem is how crustaceans utilize multiple signals obtained from the environment to track an odorant to its source by locomotion upstream. Water or air-borne odorants are detected with the aid of their antennules and in almost all crustacean species, these animals are capable of flicking their antennules. This chapter will focus on remaining problems pertaining to the fluid mechanical aspects of this phenomenon.

The antennules of aquatic and terrestrial crustaceans are their nose and are equipped with batteries of odorant-detecting sensilla. The sensilla are microscopic structures that operate at small Reynolds numbers. The Reynolds number, *Re*, is a measure of the ratio of inertial to viscous forces; it is calculated from the formula $Re = UL/v$, where U is velocity, L is diameter of the object being considered, and v is the kinematic viscosity. For large Reynolds numbers (>>1), fluid flow is turbulent and the distribution of odorants is controlled by turbulent mixing processes, while for small Reynolds numbers the flow is laminar and odorant distributions are controlled by molecular diffusion. Therefore, even though the distribution of odorants within ambient flows in oceans, streams, and the atmosphere is typically turbulent, transport on the small scale to the surfaces of chemosensory sensilla occurs under laminar conditions where slow molecular diffusion dominates. How does antennular flicking enhance odorant capture in this laminar flow environment? We herewith review what is currently understood, and what remains to be explained, about the fluid mechanics of crustacean antennular flicking.

2. Aesthetasc distribution and flicking

Odorants are scalar quantities imbedded in fluid environments. They are distributed from their sources by large-scale advective, turbulent flow as patchy filaments that vary in space, time, and density relative to the odorant source (Fig. 1). These filamentous odorant signals are sampled by the olfactory organs of the animals and may provide information about the location of the odorant source (Finelli et al. 1999; Webster et al. 2001; Grasso and Basil 2002). Microscale access of odorants to olfactory receptors in general, and to those of aquatic crustaceans in particular, is characterized by laminar flow within arrays of small specialized sensilla situated on the antennules (first antennae). These sensilla, referred to as aesthetascs in crustaceans, are blunt-tipped setae 100–1 700 µm long by 10–30 µm in diameter, depending upon the species (Snow1973; Tierney et al. 1986; Mellon et al. 1989) and occur in arrays on the ventral surface of the more lateral of two mobile flagella on each antennule (Fig. 2).

In almost all crustacean species the aesthetasc-bearing flagellum is capable of flicking behavior, by which the flagellar shaft is quickly moved directly or obliquely ventrally

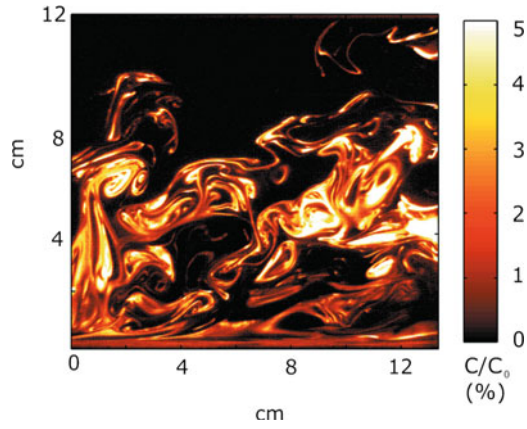


Fig. 1 Planar laser induced fluorescence image of odorant distribution 1.5 m downstream from where the source odorant was released into the flow. A 488 nm wavelength laser illuminated fluorescein dye embedded within the odorant release. Flow was created within a 25 m long laboratory flume with a mean water velocity of 7 cm/s, with movement from left to right in the image, within a total water depth of 25 cm. Concentration, *C*, is normalized by the percent of the source concentration, C_o

for a short distance and then is returned to its previous, normal posture.

Antennular flicking has attracted attention due to its obvious association with olfactory reception. This flicking is analogous to taking a "sniff" of odorant. However, the purpose of flicking has not always been met with universal agreement. This may in part be the result of anatomical differences among crustacean species with respect to the dimensions of the individual aesthetascs and, especially, with the pattern and density of their arrays on the antennular flagella. In hermit crabs and true crabs, (for example, the blue crab *Callinectes sapidus*) the aesthetascs are grouped in a high-density tuft near the tip of their lateral antennular flagellum (Fig. 2). In spiny lobsters the aesthetasc array is also dense but extends along the ventral surface of the antennule for a longer relative distance, rather than being confined to a short tuft. This lengthwise distribution of a high density array is even more extended in clawed lobsters, such as *Homarus americanus.* At the other end of the density/distribution spectrum, the array of aesthetascs in freshwater crayfish extends along the entire ventral half of the lateral antennular flagellum, and the distribution of aesthetascs along the flagellum varies from a sparse one or two per antennular annulus at the proximal end of the array to no more than four per annulus near the distal end of the array (Fig. 3).

3. Fluid mechanical consequences of flicking

In the hermit crab *Pagurus alaskensis,* the downward flick of the lateral flagellum initiates a dramatic splaying action of the dense aesthetasc tuft (Snow 1973), causing spatial separation of the distal regions of the individual sensilla, that portion of each sensillar shaft believed to be permeable to dissolved odorants (Ghiradella et al. 1968; Tierney et al. 1986). As discussed below, that physical separation and the velocity of the flick initiating the splaying action can play a critical role in the access of odorant-bearing fluid to the surface of an aesthetasc. It is not known whether this sort of splaying action is found in other species beyond crabs; it has not been reported in either spiny or clawed lobsters.

Early analyses of the fluid mechanical effects experienced by flicking antennules were undertaken by Cheer and Koehl (1987 a, b). Objects moving within a fluid volume are surrounded by a boundary layer of sheared fluid: the molecular fluid layer next to the moving object is stationary, whereas layers farther out from the object are char-

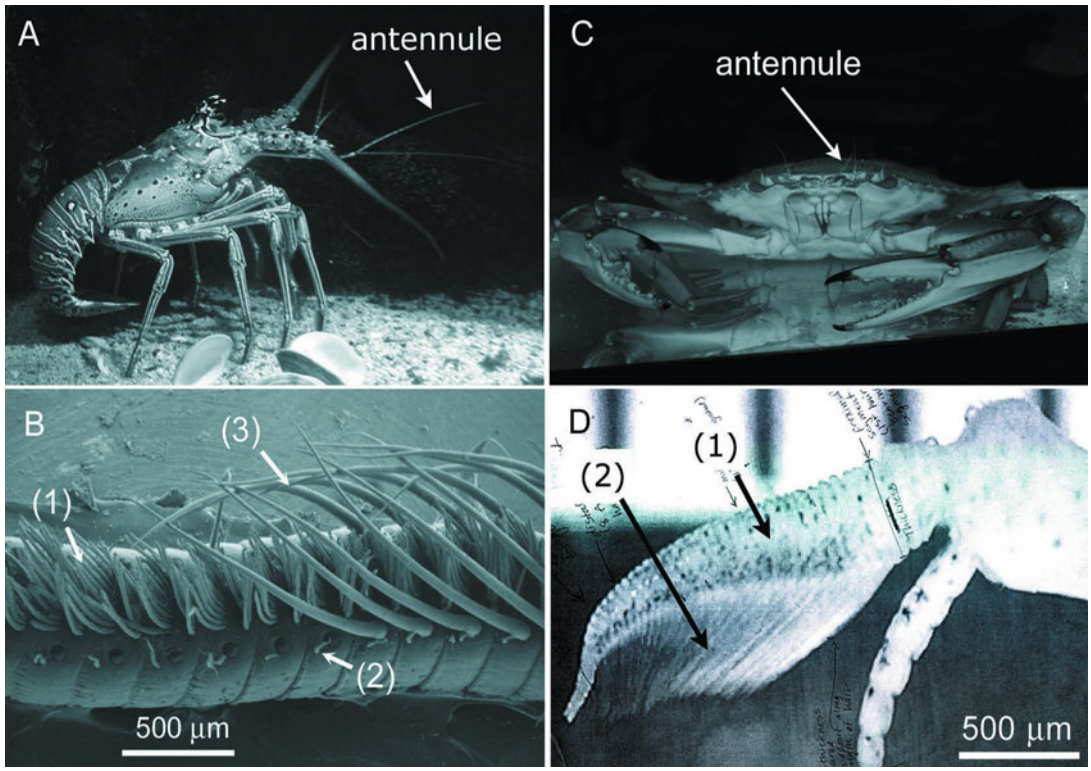


Fig. 2 **A** The spiny lobster, *Panulirus argus*, with lateral antennule labeled. **B** Scanning electron micrograph of the lateral antennule with the (1) chemosensory aesthetascs, (2) mechanosensory sensilla, and (3) guard hairs, labeled. Note that the guard hairs on the left hand side of the image have been removed to expose the aesthetascs (photo courtesy of J. A. Goldman). **C** The blue crab, *Callinectes sapidus*. **D** Antennules of *C. sapidus* with the (1) lateral antennule and (2) chemosensory aesthetascs labeled. *C. sapidus* aesthetascs splay out to increase the distance between the hairs during the flick, and clump back together during the return stroke. (photo courtesy of M. M. Martinez)

acterized by increasing velocity, eventually reaching a value identical with the bulk flow. The thickness and gradient steepness of the sheared layer next to an object depend upon the fluid's viscosity, object's size, and its velocity relative to the fluid. With very small objects, which have a low Reynolds number, or with objects moving slowly through the fluid, the boundary layer is very thick. *Re*'s of most arthropod olfactory receptor sensilla fall within the range of 10^{-4}–10 (Louden et al. 1994). When spacing between individuals of an array of tiny objects, such as aesthetascs, becomes sufficiently small, their boundary layers interfere with each other and the array acts more like a paddle than like a sieve, preventing fluid from leaking between the aesthetascs. This can present a serious impediment to the free access of odorant molecules to the surface cuticle of the aesthetascs themselves. Within limits, however, this problem may be overcome by increasing the velocity of the aesthetasc array relative to the fluid environment, thereby also decreasing the thickness of the boundary layers (Stacey et al. 2002). It has been argued that antennular flicking has evolved to sub-serve this function. For arrays of sensilla where the spacing between them is less than five times their average diameter, their leakiness is very sensitive to changes in sensillar velocity within the *Re* range of

0.1–10 (Cheer and Koehl 1987a; Koehl 1995, 2000; Koehl et al. 2001; Reidenbach and Koehl 2008). Because the relative velocity of an aesthetasc during an antennular flick depends upon its position upon the flagellar shaft, the extent and placement of the sensillar array can be important. To understand the functional significance of antennular flicking for olfactory reception, therefore, it is critical to know the range of *Re* values within which the olfactory sensilla operate, realizing that different sizes of individual aesthetascs, their spacing, placement, and the range and velocity of antennular movements play an integrated determinative role in the parameter effectiveness of the different species. The spiny lobster *Panulirus argus*, as mentioned above, possesses an aesthetasc array characterized by close spacing along a restricted region of the more distal antennular shaft. This structure has been the subject of theoretical, modeling, and experimental studies which indicate that given the mechanical properties of the aesthetascs, their position along the flagellum, and the density of their array, the downward flick of the lateral flagellum, occurring at a *Re* = 1 to 2, reduces the boundary layer of fluid around the individual aesthetascs. This allows the array to become sufficiently leaky for fluid exchange to occur during the downward flick phase of the flick-return cycle (Koehl et al. 2001; Reidenbach et al. 2008). These same studies show that, by contrast, during the slower upward return phase, occurring at a *Re* = 0.5, odorant-bearing fluid is trapped within the array, and provides time for molecular diffusion to occur across the cuticle of the individual aesthetascs, increasing the probability of odorant molecules binding with olfactory receptors within the lumen of the individual aesthetascs.

4. Comparative aspects

From comparable analyses of aesthetasc arrays characterized by very different sizes, morphologies, packing density and spatial distribution characteristics of other species, one can derive hypotheses driven by fluid mechanical considerations. Similar fluid mechanical principles that have provided insight into spiny lobster flicking should be able to be applied within this varied species-dependent context to reveal the functional significance of such diversity for the animals on which they are found. For example, hypotheses could be formed concerning the constraints that have driven the evolution of different aesthetasc densities and distributions. One such crustacean that exhibits very different aesthetasc distribution compared to that of the spiny lobster is the freshwater crayfish *Procambarus clarkii* (Fig. 3). Its aesthetasc-bearing lateral flagella of the antennules are normally held in a curved, partially-vertical orientation (Fig. 3A). The aesthetascs are arrayed ventrally along the distal one-half of each flagellum, emerging from the flagellar surface at an approximately 45° angle (Fig. 3B and C).

Figure 3C shows the distal half of the flagellum, where the aesthetasc array is characterized by pairs of (or occasionally, three) sensilla on each annulus, positioned distally. Each sensillum is approximately 100 µm long and 15–20 µm in diameter at its base, tapering to less than 10 µm at the tip. A clear discontinuity occurs in the aesthetasc structure about three fifths of the way from the base to the tip, where the cuticle becomes optically transparent (Fig. 3B). It is in this region that the aesthetascs are believed also to be transparent to odorant molecules (Ghiradella et al. 1968; Tierney et al. 1986). Therefore, the most important fluid mechanical considerations in odorant capture

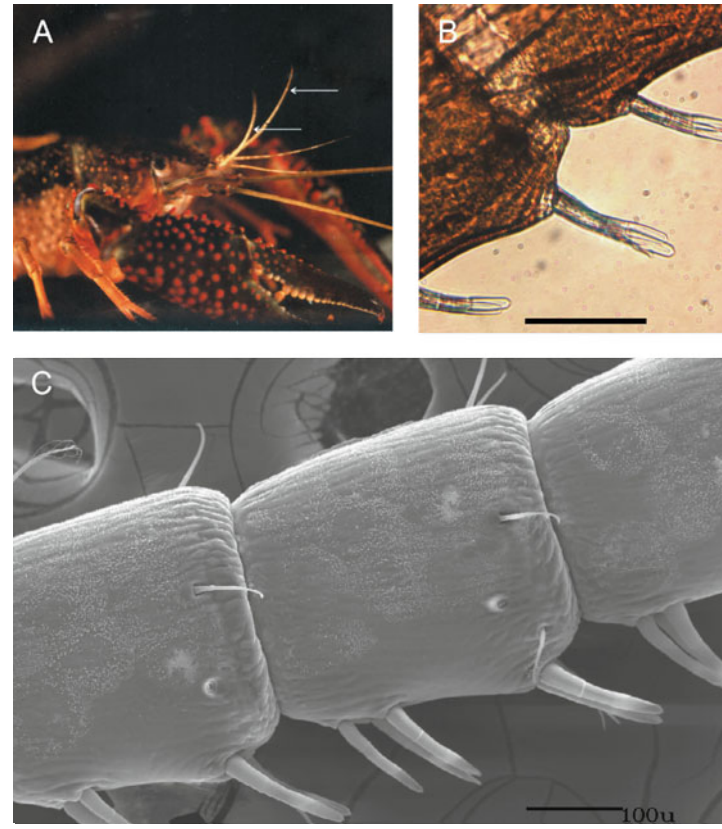


Fig. 3 **A** Front view of the crayfish *Procambarus clarkii* indicating the lateral antennular flagella (white arrows) in their usual upright resting posture. **B** Light micrograph of a living lateral flagellum and three pairs of aesthetascs on its ventral surface photographed with bright-field illumination. Note how optically transparent the distal regions are compared to the basal shafts. **C** Scanning electron micrograph of a portion of a fixed lateral flagellum showing the normal placement and spacing of aesthetascs on the flagellar annuli. Scale marks: 100 µm. **A** is reproduced with permission from Mellon and Humphrey (2007)

will be those concerning the tip region of the aesthetasc shaft. It is salient to mention that the individual aesthetascs of a pair are separated at their tip region only by about one diameter, and this has implications for leakiness of the pair to odorant-laden fluid flow.

Individuals of *P. clarkii* flick their antennules spontaneously as well as in response to hydrodynamic and odor stimuli (Mellon 1997). The video frames shown in Fig. 4 are from a series taken during a single flick cycle. Each frame represents 30 msec; Fig. 4A and 4B were taken 180 msec apart during which time the tip of the flagellum was depressed by about 6 mm*. The mean linear velocity of a point on the ventral aspect of the flagellum must therefore have been approximately 0.03 m/sec, and during the downward flick the *Re* of an aesthetasc tip region would therefore be

$$Re = (10^{-5}\,\mathrm{m} \cdot 0.03\,\mathrm{m/sec}) / 1.00 \cdot 10^{-6}\,\mathrm{m^2/sec} = 0.3$$

Although this value is less than the *Re* = 1 to 2 value calculated for spiny lobster aesthetascs during an average downward flick (Goldman and Koehl 2001), it is well within the range of values in which velocity can make a difference in the leakiness of an array of sensilla. However, as indicated by Fig. 3, side-by-side spacing of the aesthetascs in *P. clarkii* is less than one diameter, similar to the situation in another crayfish, *Orconectes propinquus* (Tierney et al. 1986, cf Fig. 1); thus, crucial questions remain concerning

* At the tip the distance traveled will be less, since the distal part of the flagellum is flexible and lags the main shaft during the initial flick (Fig. 3B).

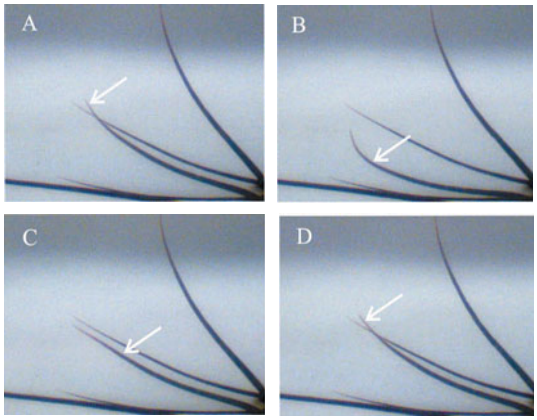


Fig. 4 Video frames from one complete flick sequence. **A** Lateral flagellum (white arrows) posture at rest, prior to downward flick. **B** Flagellum position just prior to termination of downward flick, about 180 msec after frame in **A**. **C** During the recovery phase of the flick; most of the flagellum maintains a rigid form during the upstroke. **D** Frame taken just after the flagellum arrived at the original rest posture

the effects of close inter-aesthetasc spacing and the extent to which they are sufficiently leaky to be maximally exposed to fresh odorant-bearing water during the flick cycle.

Aesthetascs occur along the ventral margin of the lateral antennular flagellum, and it is important to discuss whether fluid mechanical properties of the flagellar shaft itself influence access of odorant-laden water to the sensor array. Humphrey and Mellon (2007) mathematically modeled the flow field past the lateral flagellum during the flick cycle. The flagellum was assumed to be a rigid, tapering curved shaft hinged at the base (but see Fig. 3 for flexural properties during the flick sequence) and of cylindrical cross section (Fig. 5). During a standard downward flick and return the Reynolds number along the flagellum (Re_f) varies according to the curves in Fig. 5B, although only the region of the aesthetasc array (with a leakiness ratio: $z/L_f = 0.5 \rightarrow 1.0$, where z is any point along a flagellum of length L_f) is important for our consideration. In both movements Re_f decreases with increasing z/L_f due to a reduction in flagellum diameter and in the normal component of velocity closer to the tip. Using the Flow and Heat Transfer Solver (FAHTSO) code developed by Rosales et al. (2000; 2001) numerical calculations were made to determine the transient, developing 2-dimensional motion of water past a circular cylinder, ultimately attaining a value of $Re_f = 50$ for a downward flick or $Re_f = 30$ for an upward return trajectory. The result of these calculations are shown in Fig. 5C where U_{in}/U_o is dimensionless velocity, t^* is dimensionless time, and x_f/d_f is the dimensionless distance traveled by the approaching flow in units of the flagellum diameter. From these curves and specified boundary conditions the near-field flow streamlines around the flagellum could be calculated and visualized as in Fig. 6. An important dominant characteristic produced within the flow is the development of paired vortices in the wake of the moving flagellum (Fig. 6D). Although of little consequence during the downward flick, developing vortices within the wake of a returning flagellum could be potentially beneficial in delivering odorant-laden water to the aesthetasc array along the ventral flagellar surface. The enhanced fluid motion within the vortex can act to reduce the thickness of the boundary layer surrounding the aesthetascs, allowing for the advective transport of odors to penetrate closer to aesthetasc surfaces. This process effectively decreases the distance over which slower molecular diffusion must occur, enhancing the Pêclet number (i. e., the rate of advection of a flow to its rate of diffusion) around the aesthetasc surfaces for odorant delivery. However, the specific role of paired vortices in odorant detection needs to be critically addressed through both experimental and modeling studies. Additionally, because the distal dendrites of the olfactory neurons, with their membrane receptors, reside within the lumen of the aesthetasc, molecular

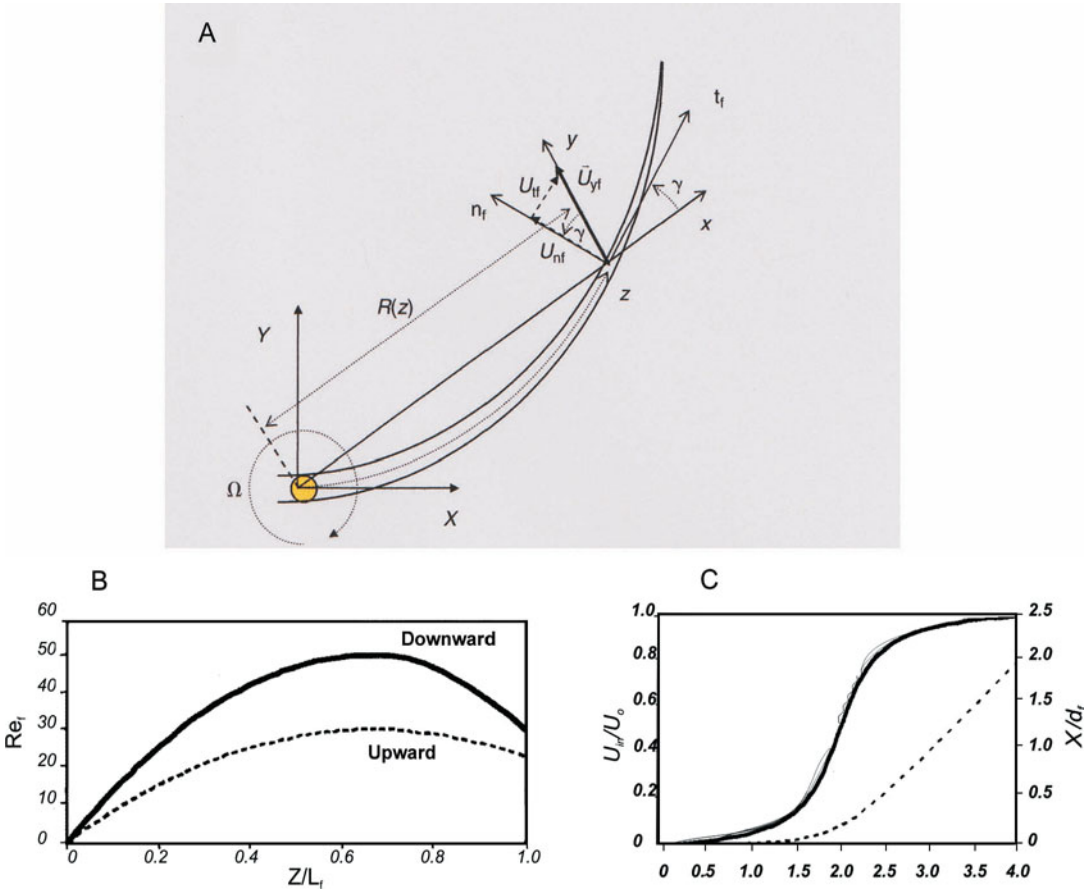


Fig. 5 **A** Schematic model of the crayfish lateral flagellum used in Humphrey and Mellon (2007). The curved, tapering shaft is assumed to be rigid and hinged (yellow circle) at the bottom during a downward flick at angular velocity Ω in relation to the pivot point in a standard X-Y coordinate system. During a flick, the approaching local fluid velocity relative to an x-y coordinate system fixed at point z and normal to the secant R(z) is U_{yf}. The approaching fluid generates components $U_{nf} = U_{yf}\cos\gamma$ and $U_{tf} = U_{yf}\sin\gamma$ normal (n_f) and tangent (t_f) to the flagellum. **B** Variation in Reynolds number Re_f along the dimensionless flagellum during a flick downward (Ω = 5.24 rad s^{-1}; heavy line) and during the return stroke (upward, Ω = 3.14 rad s^{-1}; dashed line). Decreases in flagellum diameter from base to tip and of the normal component of velocity with increasing nearness to the tip account for reduction in Re_f with increasing z/L_f. **C** Heavy line is the dimensionless velocity, U_{in}/U_o, approximating the acceleration of far-field flow approaching a flicking flagellum as a function of dimensionless time, $t^*(\sim tU_o/d_f)$. The broken line, x_f/d_f, is the dimensionless distance traveled by the approaching flow in units of flagellum diameter. For a standard flagellum, $U_o = 8.63 \times 10^{-2}$ ms^{-1} and $d_f = 5 \times 10^{-4}$ m. Reproduced with permission from Humphrey and Mellon (2007)

diffusion of odorant molecules across the sensillar cuticle must occur for effective odorant capture. While the thickness of the aesthetasc cuticle has been measured in *P. clarkii* as ~1.9 µm (Mellon et al. 1989) and in *Panulirus* as ~1.5 µm (Grünert and Ache 1988), diffusion coefficients across the aesthetasc cuticle for amino acids, which are major effective olfactory stimuli for aquatic crustaceans, remain unknown factors in this stage of the process. Ultimately, this gap in our knowledge obstructs a complete and

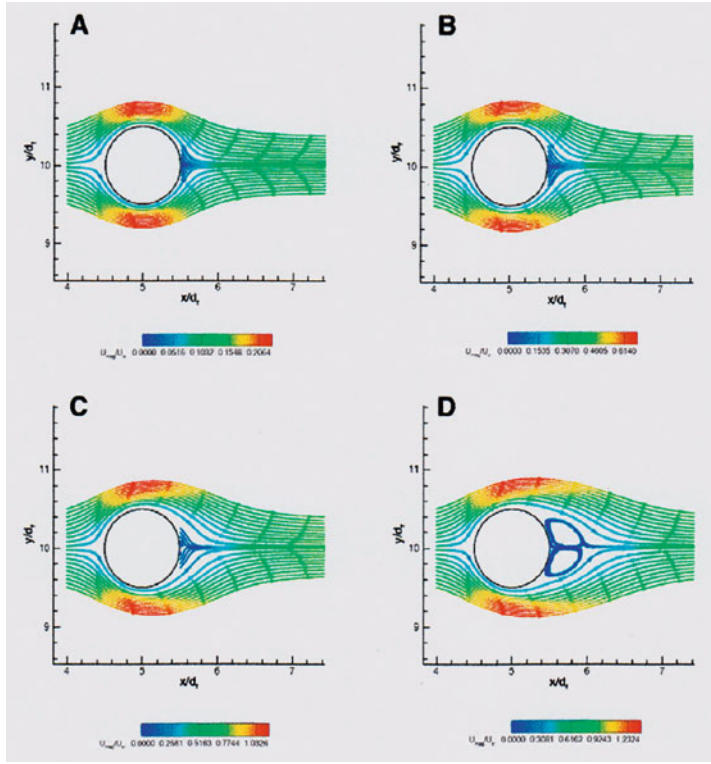


Fig. 6 Near-field flow streamlines with dimensionless velocity magnitude (color) superimposed for the 2D flow accelerating past a flagellum from left to right, according to the far field approaching velocity S-curve plotted in Fig. 5C. Results are shown **at** times t* = 1 (**A** Re_f = 2.6), t* = 2 (**B** Re_f = 25), t* = 3 (**C** Re_f = 47.1) and t* = 4 (**D** Re_f = 50). Between t* = 3 and t* = 4 the flow separates at the top and bottom of the flagellum to form a recirculating flow region containing two vortices downstream of the flagellum. Reproduced with permission from Humphrey and Mellon (2007)

comprehensive interpretation of the functional significance of antennular flicking in all crustaceans regardless of our understanding of fluid flow around the aesthetascs themselves.

5. Adaptations to terrestrial lifestyle

A number of crustaceans have adopted a secondary, obligatory terrestrial lifestyle but maintain their antennular flicking behavior. Prominent among these are members of the hermit crab family Coenobitidae, tropical crustaceans that are exclusively terrestrial as adults and which possess a highly sensitive olfactory sense (Ghiradella et al. 1968; Rittschoff and Sutherland 1986). The best known member of this family is the Caribbean land hermit crab, *Coenobita clypeatus,* individuals of which are often sold in pet stores. The largest animals of this family (in fact, the largest land-living invertebrate in the world!) are individuals of the single-species genus *Birgus latro*, native to islands of the South Pacific and Indian Oceans. Except for a brief marine larval period, these animals are entirely terrestrial and live as omnivorous scavenger/predators on remote islands. They have a well-developed sense of smell and can be attracted to suitable baits, which they track employing an antennular flicking behavior reminiscent of marine hermit crabs (Stensmyr et al. 2005). No data are available concerning the velocity of flicking antennular flagella in either *Coenobita* or *Birgus**, but if it is similar to that of *Panu-*

* Bill Hansson (personal communication) suggests that the velocity of an antennular flick in Birgus is about 1 cm/sec.

lirus (0.06 m·sec^{-1}) then the blunt, peg-like aesthetascs of *Birgus* (dimensions: 100 µm long by 40 µm in diameter) must operate at *Re* values only about 10 times smaller than those of *Panulirus,* since the kinematic viscosity of air at 25 °C is more than fifteen times that of water (15.68 × 10^{-6} m^2/sec). This means that the aesthetascs of *Birgus* operate within the range where boundary layer shedding could be achieved by antennular flicking, especially if the angular velocity turns out to be 2–3 times higher than that found in the spiny lobster.

The array of aesthetascs on the antennules of both *Coenobita* (Ghiradella et al. 1968) and *Birgus* (Stensmyr et al. 2005) is dense, and in neither case are individual sensilla separated by more than two diameters near the tip. Therefore, questions arise concerning the leakiness of the array to fresh odorant-laden fluid. Among the factors to be considered when attempting to arrive at an answer is the fact that, in contrast to the situation in aquatic environments, air-borne odorants are volatiles and as such, exist in the gaseous phase and have diffusion coefficients roughly 100 times larger than those of dissolved water-borne organic odorants. A second consideration involves the physical design of the coenobitid aesthetascs. Unlike those found on aquatic species of hermit crabs and other crustaceans like crayfish and spiny lobsters (Grünert and Ache 1988; Mellon et al. 1989), the aesthetascs are short and stubby (Fig. 7). Electron micrographic examination of the aesthetascs in *Coenobita* and *Birgus* (Ghiradella et al. 1968; Stensmyr et al. 2005, respectively) reveal them to have an asymmetrical structure: the surface cuticle of that region of the aesthetasc facing the antennule is heavy and thick, while the cuticle facing the environment is thin and crenulated, possibly as a mechanism to increase surface area. Both studies cited above have especially stressed that the short, blunt aesthetascs of *Birgus and Coenobita* are more similar to chemoreceptive hairs of insects than to the long, slender morphology of marine crustacean aesthetascs. Furthermore, the basal bodies and ciliary segments of coenobitid olfactory dendrites are, like those of insects, surrounded by a lymph space and are housed within the flagellum rather than within the sensilla, as found with aquatic crustaceans (Stensmyr et al. 2005). These differences from aquatic crustaceans suggest convergent evolutionary adaptations that address physical differences (for

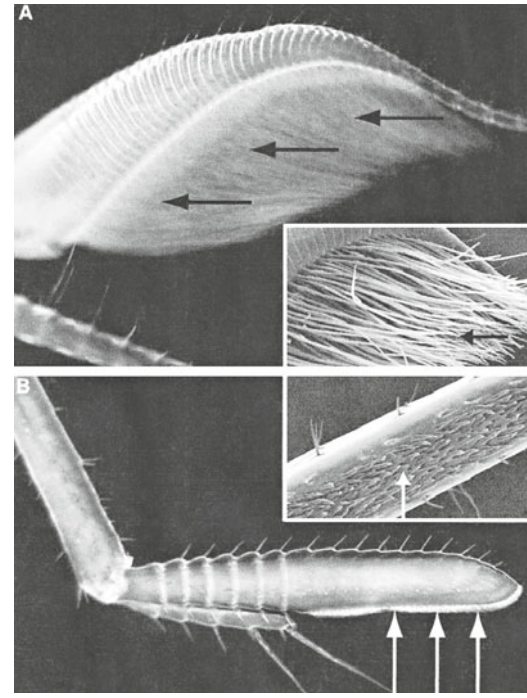


Fig. 7 **A** Distal tip of the antennule in the hermit crab *Pagurus berhardus,* a marine species. The tuft of aesthetascs (black arrows) consists of long, slender sensilla (image courtesy of S. Harzsch from Hansson et al. 2010). Inset is a scanning electron micrograph of the aesthetasc tuft, courtesy of B. Hansson and E. Sjöholm, from Hansson et al. (2010). **B** Aesthetasc array (white arrows) on the distal flagellum of the antennule in *Coenobita clypeatus* (courtesy of S. Harzsch, from Hansson et al. 2010). Inset is a SEM of aesthetascs from *Coenobita*, courtesy of B. Hansson and E. Sjöholm from Hansson et al. (2010)

example, exposure to dehydration) between the aquatic and terrestrial environments.

Conclusions

Comparative studies on the aesthetasc distribution patterns and size in a variety of different crustaceans are needed to provide a comprehensive picture of the functional significance of antennular flicking behavior to odorant capture. These studies should include experimental observations as well as modeling studies to determine the most appropriate interpretation of the morphology and anatomical distribution of aesthetascs in crustaceans with a variety of lifestyles. Of significant importance are questions about leakiness of sensillar arrays to fluid flow and the role of antennular wake vortices in abetting sensillar access to fluid flow. Adaptations to a terrestrial existence are apparent in some species as well, and these morphological differences require fluid mechanical interpretation, given the differences between the aquatic and the atmospheric environments.

Acknowledgements

The authors wish to thank Dr. J. A. C. Humphrey for helpful discussions and encouragement, Dr. W. O. Friesen for critical help with flicking video imaging, and the National Science Foundation (CBET-0933034) for partial support of the research for this chapter.

References

Cheer AYL, Koehl MAR (1987a) Paddles and rakes: Fluid flow through bristled appendages of small organisms. J Theor Biol 129: 17–39

Cheer AYL, Koehl MAR (1987b) Fluid flow through filtering appendages of insects. IMA J Math Appl Med Biol 4: 185–199

Finelli C M, Pentcheff ND, Zimmer-Faust RK, Wethey DS (1999) Odor transport in turbulent flows: Constraints on animal vegetation. Limnol Oceanogr 44: 1056–1071.

Ghiradella HT, Case JF, Cronshaw J (1968) Fine structure of the aesthetasc hairs of *Coenobita compressus* Edwards. J Morph 124: 361–385

Goldman JA, Koehl MAR (2001) Fluid dynamic design of lobster olfactory organs: High speed kinematic analysis of antennule flicking by *Panulirus argus.* Chem Senses 26: 385–398

Grasso FW, Basil JA (2002) How lobsters, crayfishes, and crabs locate sources of odor: current perspectives and future directions. Curr Opin Neurobiol 12: 721–727

Grünert U, Ache BW (1988) Ultrastructure of the aesthetasc (olfactory) sensilla of the spiny lobster *Panulirus argus*. Cell Tissue Res 251: 95–103

Hansson BS, Harzsch S, Knaden M, Stensmyr M (2010) The neural and behavioral basis of chemical communication in terrestrial crustaceans. In: Breithaupt T, Thiel M (eds) Chemical communication in crustaceans. Springer-Verlag, New York

Humphrey JAC, Mellon DeF (2007) Analytical and numerical investigation of flow past the lateral antennular flagellum of the crayfish *Procambarus clarkii.* J Exp Biol 210: 2969–2978

Koehl MAR (1995) Fluid flow through hair-bearing appendages: Feeding, smelling and swimming at low and intermediate Reynolds numbers. Symp Soc Exp Biol 49: 157–182

Koehl MAR (2000) Fluid dynamics of animal appendages that capture molecules: arthropod olfactory antennae. In: Fauci L, Gueron S (eds) Computational modeling in biological fluid dynamics. IMA volumes in Mathematics and its Applications, Vol 124.Springer-Verlag, New York, pp 97–116

Koehl MAR, Koseff JR, Crimaldi JP, McCay MG, Cooper T, Wiley MB, Moore PA (2001) Lobster sniffing: Antennule design and hydrodynamic filtering of information in an odor plume. Science 294: 1948–1951

Louden C, Best BA, Koehl MAR (1994) When does motion relative to neighboring surfaces alter the flow through arrays of hairs? J Exp Biol 193: 233–254

Mellon DeF (1997) Physiological characterization of antennular flicking reflexes in the crayfish. J Comp Physiol A 180: 553–565

Mellon DeF, Tuten HR, Redick J (1989) Distribution of radioactive leucine following uptake by olfactory sensory neurons in normal and heteromorphic crayfish antennules. J Comp Neurol 280: 645–662

Mellon DeF, Humphrey JAC (2007) Directional asymmetry in responses of local interneurons in the crayfish deutocerebrum to hydrodynamic stimulation of the lateral antennular flagellum. J Exp Biol 210: 2961–2968

Reidenbach MA, George N, Koehl MAR (2008) Antennule morphology and flicking kinematics facilitate odor sampling by the spiny lobster, *Panulirus argus*. J Exp Biol 211: 2849–2858

Rittschoff D, Sutherland JP (1986) Field studies on chemically mediated behavior in land hermit crabs (*Coenobita rugosus*): Volatile and nonvolatile odors. J Chem Ecol 12: 1273–1284

Rosales JL, Ortega A, Humphrey JAC (2000) A numerical investigation of the convective heat transfer in unsteady laminar flow past a single and tandem pair of square cylinders in a channel. Num Heat Transfer A 38: 443–465

Rosales JL, Ortega A, Humphrey JAC (2001) A numerical simulation of the convective heat transfer in confined channel flow past square cylinders: comparisons of inline and offset tandem pairs. Int J Heat Mass Transfer 44: 587–603

Snow PJ (1973) The antennular activities of the hermit crab, *Pagurus alaskensis* (Benedict) J Exp Biol 58: 745–765

Stacey MT, Mead KS, Koehl MAR (2002) Molecule capture by olfactory antennules: Mantis shrimp. J Math Biol 44: 1–30

Stensmyr MC, Erland S, Hallberg E, Wallén R, Greenaway P, Hansson BS (2005) Insect-like olfactory adaptations in the terrestrial giant robber crab. Curr Biol 15: 116–121

Tierney A-J, Thompson CS, Dunham DW (1986) Fine structure of aesthetasc chemoreceptors in the crayfish *Orconectes propinquus.* J Can Zool 64: 392–399

Webster DR, Weissburg MJ (2001) Chemosensory guidance cues in a turbulent chemical odor plume. Limnol Oceanogr 46: 1034–1047

Stagnation point flow analysis of odorant detection by permeable moth antennae

12

Joseph A. C. Humphrey†, Hossein Haj-Hariri

Contents

Abstract

A mathematical model is presented for the transport of odor stimuli to the chemoreceptor sensilla of moth antennae. The odorant is in the form of threadlike filaments composing a meandering plume. The classical solution for two-dimensional stagnation-point flow (also called Hiemenz flow) is extended to include a permeable or 'leaky' surface. The surface of an antenna is represented as a longitudinal section of the surface of a cylinder of arc length $s = \theta a$ where a is the radius of the cylinder and θ is the angle subtended by the arc. In order to simulate leakage to the flow approaching normal to this permeable surface, the flow around a slightly smaller cylinder of radius $a - \varepsilon$ with $\varepsilon \ll a$ is considered first. The flow in the vicinity of the stagnation point is then determined. If $\varepsilon = 0$, non-permeable stagnation point flow conditions are recovered, otherwise stagnation point flow conditions with radial leakage into the surface are obtained. The leakage is found to be proportional to ε/a, with a $\cos\theta$ dependence which provides a good approximation to a constant leakage in the vicinity of the stagnation point. The solution to the ordinary differential equation describing leaky Hiemenz flow depends on the Reynolds number Re_d ($\equiv 2av_\infty/\nu_f$) based on the cylinder diameter, the approaching fluid velocity and the kinematic viscosity, and ϕ ($\equiv v_o/v_\infty$), the leakage parameter with v_o denoting the speed of the flow passing through the permeable surface. With the velocity field known, the distributions of odorant species concentration in the flow (i) approaching an antenna, (ii) on the surface of an antenna, and (iii) within the chemosensilla attached to the antenna, are obtained from respective conservation equations and boundary conditions. Numerical results are presented illustrating the effects of relevant parameters such as Re_d, ϕ, and Sc ($\equiv \nu_f/D$, the Schmidt number), where D is the diffusion coefficient of the odorant species in air.

Hossein Haj-Hariri
University of Virginia, Department of Mechanical and Aerospace Engineering
122 Engineer's Way
Charlottesville, VA 22904-4746, USA
e-mail: hh2b@virginia.edu

1. Introduction

1.1 The problem of interest

The behavior of many land- and water-based animals is affected by the ability of their olfactory organs to detect certain chemical substances mixed with, dissolved in, or suspended in the surrounding fluid medium, often in very small amounts. These chemicals are also referred to as 'odorants' and may consist of a single compound or a blend of compounds in which case the composition of the blend can be very important. In the case of insects, for example: air-borne odorants from host plants or animals mediate the availability of food, oviposition sites, egg-laying and other functions; pheromones play major roles as chemical signals in intraspecific communication such as mate-finding, aggregation and alarm; and some species have developed defensive chemical compounds that repel other insect and vertebrate predators.

Odorant generation, dispersion and detection, and olfaction-mediated orientation, have been subjects of longstanding interest among biologists and, as a consequence, the literature for insects alone is vast (Wright 1958, 1964, 1982; Bossert and Wilson 1963; Schneider 1969; Kaissling 1971, 1985, 1986, 1997, 2004; Bell and Cardé 1984; Cardé 1984; Cardé and Baker 1984; Elkington and Cardé 1984; Morita and Shiraishi 1984; Mustaparta 1984; Kennedy 1986; Payne et al. 1986; Murlis et al. 1992; Vickers and Baker 1992, 1994; Shuranova and Burmistrov 1996; Hansson 1999; Barth and Schmid 2001; Barth et al. 2002; Loudon and Davis 2005). Many of these references consist of fundamental studies seeking to understand and explain the complex phenomena connecting odor generation, dispersion, and detection with observed animal behavior. The majority of the studies is experimental in nature, ranging from making field observations of animal behavior in response to the spatio-temporal characteristics of odor plumes evolving under natural conditions, to performing intra- or extra-cellular electrophysiological recordings of the action potentials fired by olfactory receptor neurons under laboratory-controlled stimulus odorant concentration and flow conditions. There are few studies however seeking to model and calculate the multiscale range of the physical-chemical phenomena observed. Aside from their intrinsic fundamental value, given the importance of some of the more applied aspects of insect olfaction aimed at, for example, using semiochemicals for pest management and control (Karg and Suckling 1999), predictive models have the capacity to provide very useful information in fractions of the time and cost of corresponding experiments.

In this study we develop and apply a multi-scale theoretical framework for modeling the physics of odorant detection by permeable moth antennae in air for biologically relevant conditions. While emphasis is placed on the antennae of moths, within the constraints imposed the model can be applied to the antennae of other insects. The interrelated objectives are three: 1) to simulate the essential physics underpinning odorant transport and capture by the antennae; 2) where possible, to compare the performance of the model with respect to the relevant biological data available; 3) to use the model to explore a meaningful range of biological and physical-chemical conditions affecting odorant detection.

1.1.1 Odorant-elicited anemotactic behavior of moths in flight

Of special interest is the detection of pheromones and plant odorants by moths flying in air. In particular, the detection by male

moths of intermittent filaments or patches of pheromones emitted by calling female moths leads to positive optomotor anemotactic behavior which, on average, guides the males upwind towards the odorant source (Kaissling 1997). The physics, physiology, and anemotactic behavior associated with pheromone emission by female moths and its detection by male moths have been investigated extensively by, among others, Kaissling (1971, 1985, 1997, 2004), Percy and Weatherston (1974), Bjostad et al. (1980), Kennedy (1983, 1986), Baker (1986), David (1986), Preiss and Kramer (1986), Schneider (1986), Vickers and Baker (1992, 1994) and Mafra-Neto and Cardé (1994). Briefly, the females of many moth species pulse (protrude and retract) their pheromone-emitting gland with a periodicity of 1 second or longer. In the case of 1-day old female gypsy moths (*Lymantria dispar*), the mean pulse period is about 3.2 seconds, with a range of 0.5 to 50 seconds (Cardé et al. 1984). Eversion of the female gland exposes its cuticle, under which lie the secretory cells (Percy and Weatherston 1974). The absence in most female moths of pore canals or ducts connecting the secretory cells to the cuticle has led to the proposition that a diffusion gradient is maintained across the cuticle by the evaporation of pheromone from the gland surface (Percy and Weatherston 1974). In this way, and depending on the prevailing air flow conditions in the vicinity of the female moth and past its abdominal tip, periodic filaments or patches of pheromone vapor result. Thus, the amount of pheromone vapor emitted during a release period is a function of the surface area of the exposed gland, the pheromone concentration on the gland surface, and the flow conditions past the gland surface. A notable exception has been observed in calling females of the arctiid species, *Pyrrharctia isabella*, and its relatives, in which the periodic emission of pheromone from invaginated tubular glands occurs as a stream of liquid droplets ranging in size between 5 and 20 μm, approximately (Krasnoff and Roelofs 1988).

The pheromone-laden odorant generally consists of a blend of chemical compounds in fairly specific concentration ratios. It is gradually dispersed and diluted in the air medium by convective/diffusive transport downwind from the source. If, as generally is the case, the air motion is turbulent, this leads to the formation of highly tortuous, three-dimensional odorant filaments. However, because the initial diameter of a filament is of order or smaller than the smallest viscous-dominated (Kolmogorov) scales of motion, viscous (molecular) diffusion initially dominates filament spreading and, as a consequence, the identity of the filament, even though tortuous, can be preserved over relatively long distances.

When detected even in minute quantities by the hair-like chemoreceptor sensilla on the antennae of a male moth, the pheromone triggers an optomotor anemotactic behavioral response (Kaissling1997). The animal then orients to fly upwind, sometimes directly and sometimes following a zig-zag pattern, while maintaining visual contact with the ground. As long as odorant continues to be detected intermittently within a frequency range spanning a few Hertz the moth maintains a general upwind course. If the odorant ceases to be detected, the moth eventually abandons the upwind-directed course to fly laterally, in search of new odorant filaments or patches. This behavior is called “casting,” and soon after new odor is detected the moth resumes anemotactic orientation. This overall behavior continues until the moth loses all contact with the odor or makes visual contact with a female.

1.1.2 Morphologies of moth antennae and their chemosensilla

Studies of the morphology, physiology and chemoreception of insect antennae and their sensilla have been performed by Schneider (1964), Callahan (1975), Berg and Purcell (1977), Berg (1983), Futrelle (1984), Morita and Shiraishi (1985), Zacharuk (1985), Payne et al. (1986), Laue and Steinbrecht (1997), Kaissling (1998, 2001), Steinbrecht (1997, 1999), and Hansson (1999). Scheer (2003) presents a beautiful collection of color plates showing especially clearly some of the possible variations in morphology of male moth antennae. The antennae of some moths such as *Laothoe juglandis*, *Manduca sexta* and *Smerinthus jamaicensis* are threadlike, consisting of a main, cylindrically-shaped shaft, with length to diameter ratio $L/d >> 1$, from which fairly short branches protrude which support the fine chemosensing hairs (sensilla). The branches in the antennae of other moths such as *Antheraea polyphemus*, *Hyalophora cecropia*, and *Actias luna* are much longer giving the antenna a comb-like (or pectinate) appearance. For the *Lepidoptera*, of which moths constitute the majority of species, the antennae shape may be: a) filiform or thread-like, with no thickening of the antenna shaft or branches extending from the antenna shaft; b) clubbed, as the result of a thickening of the antenna shaft towards its tip; c) pectinate or comb-like, with branches extending from one side of the antenna shaft; d) bipectinate or feather-like, with branches extending from both sides of the antenna shaft, with one pair of extensions at each junction; e) quadripectinate or doubly pectinate, like the bipectinate but with two pairs of branches at each junction. Whereas a filiform antenna resembles a long and relatively straight cylinder, the sensilla-supporting branches of pectinate, bipectinate and quadripectinate antenna define extended permeable surfaces that may be flat or curved.

The orientation of a moth's antennae relative to the prevailing direction of the air stream passing over the animal's body during flight is important but, unfortunately, the information available is sparse. The problem is complicated by the fact that the antenna orientation that favors maximum interception of odorant species is normal to the prevailing air velocity and this maximizes drag. The speed of the flow past the antenna of a gypsy moth (*Limantria dispar*) can be as high as 1.5 m/s (Willis et al. 1994). For bipectinate and quadripectinate antennae with the open (concave) side facing the flow, the drag forces on the branches and the hair-like sensilla they support may work to splay the antenna open so that it is less concavely curved than at rest and more porous.

In the case of filiform-shaped antennae, the sensilla extend from the hair shaft giving the antenna a bottle-brush-like appearance under magnification. In the case of antennae with branches, the sensilla extend from the branches into the spaces between branches. The analysis of this study applies to both filiform-shaped and flat or mildly curved (whether convexly or concavely) pectinate, bi-pectinate and quadripectinate antennae facing the approaching odorant-laden flow.

1.1.3 The multi-scale nature of odorant detection

The process of odorant detection is multi-scale in nature involving: (i) a macroscale plume of diameter $d \cong 1–10$ m and length $L \cong 10–1\,000$ m containing dispersed, tortuous odor filaments ($d \cong 1–10$ mm); (ii) the mesoscale antennae ($d \cong L \cong 1–10$ mm); (iii) the microscale sensilla ($d \cong 1–5$ µm, $L \cong 10–10^3$ µm) on the antennae; and, (iv) the chemo-receptor sites (pore/tubule units with $d \cong 10$ nm and $L \cong 10^2$ nm) located on the sensilla. All of these are only approximate diameter (d) and length (L) scales. While

much is known about the characteristics of plumes, antennae, sensilla and the receptor sites in their respective length scale ranges, comparatively less has been said about the overlap between scales where the odor 'information transfer' affecting behavior takes place. This is particularly true for the overlap between the mesoscale and microscale ranges, and extending into the nanoscale range. The characteristic velocities or rates associated with the physical-chemical phenomena taking place at these length scales yield corresponding time scales spanning 10^{-4} to 10 s, approximately.

Elkington et al. (1984), Murlis and Jones (1981), Murlis (1986) and Murlis et al. (1992) have reviewed the literature on odor plumes and how insects use them. Important has been the recognition that, because odorant species concentration varies randomly in both space and time, time-averaged field approaches for describing odor plumes do not properly characterize the physical situation since moths do not track odorant gradients. Instead, their self-directed anemotatic behavior is triggered by the intermittent detection of discrete odor filaments or patches (Kennedy 1986; Vickers and Baker 1992, 1994; Mafra-Neto and Cardé 1994).

Analyses of the flows past antennae and their sensilla have been performed by Adam and Delbrück (1968), Murray (1977), Vogel (1983), Loudon and Koehl (2000), Loudon and Davis (2005). Whereas an odorant-laden flow must pass around a filiform-shaped antenna, some of the flow will pass between (or leak through) the branches of pectinate, bipectinate or quadripectinate antennae. The amount of flow passing through one of the latter antennae is a function of antenna permeability which, in turn, depends on the number and distribution of the branches as well as the number and distribution of the extremely fine hair-like chemo-sensilla supported by the branches. Vogel (1983) finds that the speed of the air flowing through the antenna of the moth *Actias luna* ranges between 0.05 and 0.18 m/s for approaching air speeds ranging between 0.5 and 2.5 m/s. As odor flows over or through an antenna, some fraction of the odorant molecules is adsorbed on the surfaces of the sensilla supported by the antenna. Whether via a direct hit or 2 D random motion over the sensillum surface, some of these molecules find their way to nanometer-sized pores that connect to tubules filled with lipophilic materials that allow odor chemicals a diffusion-mediated passage along the tubules. In the case of pheromones, pheromone-binding proteins are responsible for transporting the odorant from the end of the tubule, through the sensillum lymph towards a receptor site on the membrane of one of the dendrites within the sensillum lymph (Kaissling 1998, 2001, 2004, 2009). The activation of a receptor site opens an ion-conducting channel on the dendrite membrane which eventually results in the firing of action potentials.

2. The physical-mathematical model: flow and concentration fields

2.1 Prelimary considerations

In principle, detailed distributions of the velocity and species concentration fields for the odorant-laden flow past an antenna and its sensilla can be obtained by numerically solving three-dimensional, unsteady forms of the continuity, momentum and species transport equations. However, to retain physical accuracy, such an approach requires the application of computationally costly refined grids, preferably using adapted coordinates, to resolve the effects of the complicated antenna geometry on the flow.

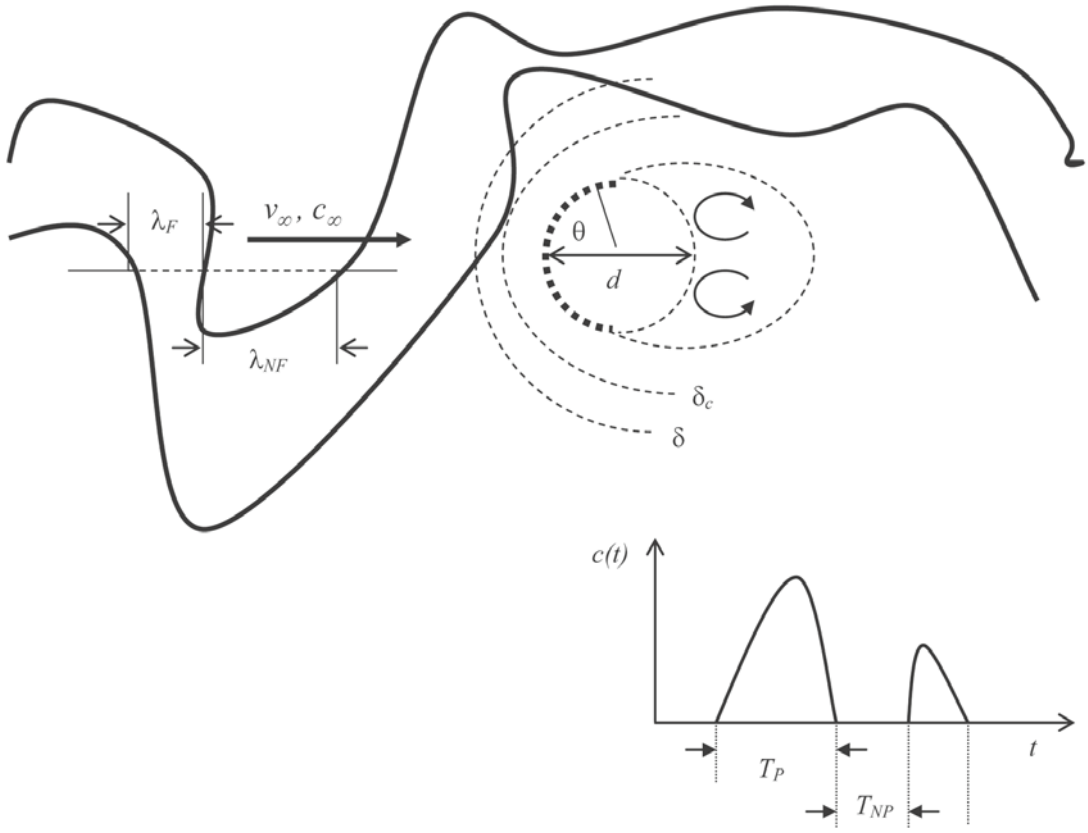


Fig. 1 Schematic showing a tortuous time-dependent odorant filament of characteristic width λ_F and intermittency λ_{NF} approaching an idealized permeable moth antenna at far field velocity v_∞ and concentration c_∞. The insert shows the variation of species concentration $c(t)$ at the edge of the antenna concentration boundary layer, δ_c as well as the time scales $T_P = \lambda_F/v_\infty$ and $T_{NP} = \lambda_{NF}/v_\infty$ associated with the filament width and its intermittency. The antenna is approximated by an arc-shaped section of the curved surface of a permeable cylinder of diameter d. For $Re_d = dv_\infty/v_f \geq 10$ and small permeabilities the flow can separate at the edges of the antenna causing mild recirculation in its wake, but this does not affect the stagnation point flow analysis. The dimensionless velocity boundary layer thickness is given by $\delta/d \infty Re_d^{-1/2}$ and the dimensionless concentration boundary layer thickness by $\delta_c/\delta \infty Sc^{-0.41}$. The figure shows the convex antenna surface facing the oncoming flow. However, the analysis also applies to concave antenna surfaces facing the oncoming flow provided stagnation point flow conditions apply (White 1974, Panton 1996)

In addition, the lack of accurate boundary condition information and values for some of the parameters needed to computationally close the odorant detection problem makes this approach undesirable.

In this study we develop an analytical solution approach that is underpinned by a number of assumptions and approximations which, while rendering the physical-mathematical model mainly qualitative, retains all the essential physics and results in the fast, effective and inexpensive calculation of parameter dependencies. Indeed, what is sacrificed in terms of predictive quantitative accuracy is gained in terms of fundamental qualitative understanding within the context of a generally applicable and computationally simple odorant transport model.

Numerical calculation accuracy is not a limitation in this alternative approach.

The analysis is focused on the antennae of moths. With reference to Fig. 1, the model is two-dimensional, implying that the characteristic length of an antenna is significantly longer than its characteristic equivalent diameter. This is not an unduly restrictive assumption within the context of the stagnation point flow analysis developed below. The antenna is assumed to correspond to a curved arc-shaped section of the surface of a long, straight cylinder of effective radius $a = d/2$. The internal polar angle, θ, subtended by the concave side of this surface is generally much smaller than 360° but is equal to 360° for the circular-shaped antennae of some moths[1].

The antenna surface is assumed to be oriented normal to an odorant-laden flow approaching at far-field velocity v_∞ and with concentration $c_\infty(t)$. The flow is taken to be quasi-steady over the characteristic time required for odorant detection, but the convected odor concentration field is a function of time and space. The flow over the antenna generates a velocity boundary layer of thickness δ and a concentration boundary layer of thickness δ_c such that, in air, $\delta_c/\delta < 1$ (Bird et al. 1960). The time of exposure of the antenna to an odor filament is T_P and the time between exposures is T_{NP}. If branched, as in the case of some moths, the antennae are permeable to the approaching flow. Following Vogel (1983), a permeability or leakage parameter, $\phi \equiv v_o/v_\infty$, is defined for the antenna where v_o is the velocity of the flow passing through the antenna, essentially normal to a fictitious "continuous effective surface[2]" defined by the primary branch, and the secondary branches supporting the sensilla. Analysis is restricted to flows for which $v_o/v_\infty << 1$, and spatially to a small region to either side of the stagnation point on the antenna. Again, these are not serious constraints as the resulting model applies to flows with $v_o/v_\infty \leq 0.1$, and to curved sections of the cylinder surface of arc length $s/a \leq 0.1$ corresponding to an angle $2\theta = 6°$ about the stagnation point, approximately, both of which represent biological meaningful conditions.

2.2 The flow field

In order to simulate possible antenna leakage to the approaching flow we extend the classical 2D stagnation point flow solution or Hiemenz[2] flow (Milne-Thomson 1968, Schlichting 1968, Batchelor 1967, White 1974, Panton 1996) as follows. We consider first the approaching inviscid, irrotational,

[1] Fig. 1 shows the convex antenna surface facing the oncoming flow. However, the analysis also applies to concave antenna surfaces facing the oncoming flow provided stagnation point flow conditions apply (White 1974; Panton 1996).

[2] The 2D stagnation point flow analysis past an impermeable cylinder (Hiemenz flow) consists of two steps: (i) first, the approaching inviscid (zero viscosity), irrotational (zero vorticity) flow approaching the cylinder is found from its complex potential; the complex potential technique is nicely explained in, for example, Milne-Thomson 1968; (ii) then from this a viscous steady state flow is determined that has the same general structure very near to the cylinder surface as explained in, for example, Panton 1996. Thus, for the antenna flow field analysis (but not the odorant field analysis) it is necessary to define a fictitious "continuous effective surface" encompassing the antenna primary branch, the secondary branches and the sensilla. This can be done because even though permeable antennae are not shaped as true cylinders at a distance they appear so to the approaching flow and, in the vicinity of the stagnation point, flow leakage through the antenna surface is small enough for a perturbation of the Hiemenz analysis to apply.

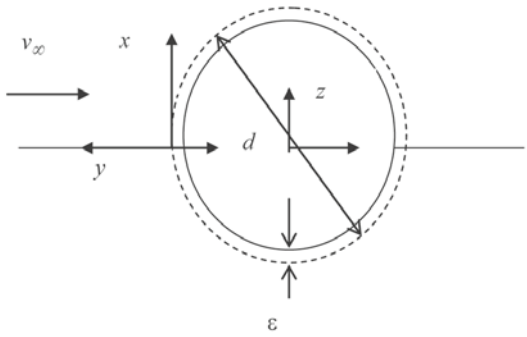


Fig. 2 Geometry for the analysis of the flow around a leaky cylinder of diameter d = 2 approximating a moth antenna. The coordinates x and y are external to the cylinder, with y pointing counter to the approaching flow. The quantity z is a complex number representing the coordinates internal to the cylinder. The locations z = –a (the real part of z) and x = 0, y = 0 coincide at the flow stagnation point

potential flow around a cylinder of radius $a - \varepsilon$ with $\varepsilon \ll a$ as depicted in Fig. 2[3].

We then determine the viscous flow in the vicinity of the point $z = -a$ ($-a$ is the real part of z, a complex number) and, if $\varepsilon = 0$, we recover the stagnation point flow conditions at $z = -a$, otherwise we obtain stagnation point flow conditions with radial leakage into the cylinder. With reference to Eq. (3) below, this leakage is found to be proportional to ε/a and to have a cos θ dependence. The cos θ dependence provides a good approximation to a constant leakage in the vicinity of the stagnation point.

2.2.1 Ideal flow around an impermeable cylinder of radius a – ε

The complex potential of the flow, *F*, and the corresponding complex velocity, *W*, for the ideal, inviscid, irrotational, potential flow around an impermeable circular cylinder of radius $a - \varepsilon$ ($\varepsilon \ll a$) with approaching velocity v_∞ along the real $-z$ axis are given by Milne-Thomson (1968) as

$$F(z) = v_\infty \left(z + \frac{(a-\varepsilon)^2}{z} \right) \tag{1}$$

and

$$W(z) = v_\infty \left(1 - \frac{(a-\varepsilon)^2}{z^2} \right), \tag{2}$$

where *z* is the complex variable with origin at the center of the cylinder. Because the flow region of interest lies around the point $z = -a$, the change of variables, $\zeta = z + a$, suggests itself. From this it follows that

$$W(\zeta) = v_\infty - \frac{(1-\varepsilon/a)^2 v_\infty}{\left(1 - \frac{\zeta}{a}\right)^2}. \tag{3}$$

Expanding the above expression for the limit of small ε/a and small $|\zeta|/a$ yields

$$W(\zeta) = v_\infty \left[1 - \left(1 - \frac{2\varepsilon}{a} \right) \left(1 + \frac{2\zeta}{a} \right) \right] = \frac{2v_\infty}{a} \left[\varepsilon - \zeta + \frac{2\varepsilon\zeta}{a} \right] \tag{4}$$

and, since $W(\zeta) = U - iV$, it follows that

$$U = \frac{2v_\infty}{a} \left[\varepsilon - \left(1 - \frac{2\varepsilon}{a} \right) \Re(\zeta) \right] \tag{5-a}$$

and

$$V = \frac{2v_\infty}{a} \left(1 - \frac{2\varepsilon}{a} \right) \Im(\zeta) \tag{5-b}$$

where $\Re(\zeta)$ and $\Im(\zeta)$ are the real and imaginary components of ζ.

2.2.2 Local ideal leaky flow expressed in terms of the x-y coordinate system

In the remainder of this analysis, to be consistent with the assumption that $\varepsilon/a \ll 1$,

[3] By specifying $a - \varepsilon$ with $\varepsilon \ll a$ an analysis can be performed that is not encumbered with intractable higher order terms. We find later that $\varepsilon/a < 0.1$ is sufficient for practical purposes.

all terms quadratic and higher in ε/a are neglected. In order to transfer the above results from the local ζ plane into the real x-y coordinate system shown in Fig. 2, we note the following relations:

$$y = -\Re(\zeta),\ x = \Im(\zeta),\ u_x = V, v_y = -U\ . \quad (6)$$

Thus, the expressions for the velocity components in the x-y coordinate system are

$$u_x = \frac{2v^*}{a}x \quad (7a)$$

and

$$v_y = -\frac{2v^*}{a}(\varepsilon + y) \quad (7b)$$

where

$$v^* = v_\infty\left(1 - \frac{2\varepsilon}{a}\right)\ . \quad (8)$$

The quantity ε can be defined in terms of physical parameters of the problem. For this, we note that the normal velocity on the cylinder surface, $y = 0$, is $v_y = -v_o$, where v_o ($<< v_\infty$) is prescribed. Coupling this information with the form of the mathematical expression for v_y as given by Eq. (7b) yields

$$\varepsilon = \frac{a}{2}\frac{v_o}{v_\infty} \quad (9)$$

thus establishing the physical value for the mathematical parameter ε. Inserting Eq. (9) into Eq. (8) gives

$$v^* = v_\infty - v_o \quad (10)$$

and it follows that the physical forms of the x- and y-components of velocity are

$$u_x = \frac{2\left(v_\infty - v_o\right)}{a}x \quad (11a)$$

and

$$v_y = -\frac{2(v_\infty - v_o)}{a}\left(\frac{av_o}{2v_\infty} + y\right). \quad (11b)$$

2.2.3 The leaky Hiemenz flow solution

The expressions for u_x and v_y given by Eq. (11a) and (11b) are taken as corresponding to the free-stream velocity components around a leaky cylinder. Because of the no-slip boundary condition at the cylinder surface, the velocity varies sharply in a thin boundary layer of thickness δ. The boundary layer analysis in this section parallels the Hiemenz solution for two-dimensional stagnation point flow given in, for example, White (1974) and Panton (1996). We extend that solution to allow flow through a permeable cylindrical surface, corresponding to a leaky moth antenna.

At the cylinder surface Eq. (11a) and (11b) must satisfy no-slip, permeable surface boundary conditions given by

$$u_x(x,0) = 0, \quad v_y(x,0) = v_o \quad . \quad (12)$$

The boundary layer thickness, δ, to be determined as a part of the solution, is used to nondimensionalize the surface-normal y coordinate to give $\eta \equiv y/\delta$. Also, to simplify the notation, henceforth we use u and v in place of u_x and v_y to denote the x- and y-components of velocity. Assuming that a similarity solution exists, the following form for the x-component of velocity inside the boundary layer suggests itself

$$u = \frac{2v^*}{a}f'(\eta)x \quad (13a)$$

or, equivalently,

$$u = \frac{2(v_\infty - v_o)}{a}f'(\eta)x\ . \quad (13b)$$

Compatibility with the surface boundary condition, $u(x,0) = 0$, requires $f'(0) = 0$. The boundary condition at a large distance beyond the edge of the boundary layer is $f'(\eta \to \infty) = 1$.

The similarity form for v is determined through an application of the continuity equation

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}=0 \tag{14}$$

from which it follows that

$$\frac{2v^*}{a}f'(\eta)+\frac{1}{\delta}\frac{\partial v}{\partial \eta}=0\,. \tag{15}$$

Integration of this equation gives

$$v=-2v^*\frac{\delta}{a}f(\eta)+C \tag{16}$$

where the constant C is determined from the permeable surface boundary condition, $v(x,0)=-v_o$. Thus, we find

$$v=-2v^*\frac{\delta}{a}f(\eta)-v_o \tag{17a}$$

with $f(0)=0$. However, the expected smallness of δ/a results in the following expression for the y-component of velocity

$$v=-2v_\infty\frac{\delta}{a}f(\eta)-v_o\,. \tag{17b}$$

As in the Hiemenz flow solution, the equation for the y-component of momentum simply yields that pressure p is independent of the coordinate direction y. The equation for the x-component of momentum is the essential equation of interest. However, this equation contains the streamwise pressure gradient, dp/dx, an unknown quantity which is the driving term in the equation. The pressure gradient is determined from an application of the Bernoulli equation to the free-stream to obtain

$$\frac{1}{\rho}\frac{dp}{dx}=-\left(\frac{2v^*}{a}\right)^2 x \tag{18}$$

Substituting Eq. (13a), (17a) and (18) into the equation for the x-component of momentum and neglecting streamwise diffusion, yields

$$\left(\frac{2v^*}{a}\right)^2 xf'^2-(2v^*\frac{\delta}{a}f+v_o)\frac{2v^*}{a}\frac{x}{\delta}f''=\left(\frac{2v^*}{a}\right)^2 x+v_f\frac{2v^*}{a}\frac{x}{\delta^2}f''' \tag{19a}$$

or, equivalently,

$$f'^2-\left(f+\frac{v_o}{v^*}\frac{a}{2\delta}\right)f''=1+\frac{v_f a}{2v^*\delta^2}f'''\,, \tag{19b}$$

where the dependence of f on η is implied in the notation and v_f is the fluid kinematic viscosity.

Based on Eq. (19b) we now define the boundary layer thickness, δ, as

$$\delta=\sqrt{\frac{v_f a}{2v_\infty}}\,. \tag{20}$$

Initially, it might seem from the form of the coefficient multiplying f''' in Eq. (19b) that the shifted velocity, v^*, should be used instead of the approaching velocity, v_∞, in the definition of δ. However, such a definition complicates the comparison of results of simulations using different permeable surface through-flow velocities, v_o, because of the intricate dependence of δ on v_o. Recalling that we have assumed $v_o/v_\infty \ll 1$, the final x-momentum equation reads

$$f'^2-\left(f+\frac{\phi}{2}\mathrm{Re}_d^{1/2}\right)f''=1+(1+\phi)f''' \tag{21}$$

where Re_d ($\equiv dv_\infty/v_f$) is the Reynolds number based on the cylinder diameter and ϕ ($\equiv v_o/v_\infty$) is the leakage parameter. This equation can be solved numerically using a Runge-Kutta scheme subject to the boundary conditions

$$f(0)=0,\ f'(0)=0,\ f'(\eta\to\infty)=1\,. \tag{22}$$

We refer to the solution of Eq. (21) and its associated boundary conditions as the "leaky Hiemenz flow" (LHF) solution. It is clear from the form of this equation that the flow field depends only on Re_d and ϕ. Clearly, when $\phi = 0$ the standard Hiemenz flow solution is recovered. In the following section the more general LHF base flow is incorporated into the equation describing the transport of a chemical species towards the cylinder.

2.3 The concentration field

The equation describing the transport of a chemical species of concentration $c(x, y, t)$ in the flow approaching the cylinder is given by

$$\frac{\partial c}{\partial t} + u\frac{\partial c}{\partial x} + v\frac{\partial c}{\partial y} = D\left(\frac{\partial^2 c}{\partial x^2} + \frac{\partial^2 c}{\partial y^2}\right). \qquad (23)$$

This equation can be written in nondimensional form by defining a characteristic time scale, $t_{sc} = a/2v_\infty$ and scaling variables as follows: t with t_{sc}, x with a, y with δ, u with x/t_{sc}, v with $(v_f/t_{sc})^{1/2}$, and c with c_{ref}, where c_{ref} can be taken as the maximum species concentration approaching the cylinder. The species transport equation is then

$$\frac{\partial C}{\partial \tau} + \alpha f'\frac{\partial C}{\partial \alpha} - \left[f + \frac{\phi}{2}\mathrm{Re}_d^{1/2}\right]\frac{\partial C}{\partial \eta} = \frac{1}{Sc}\left[\left(\frac{\delta^2}{a^2}\right)\frac{\partial^2 C}{\partial \alpha^2} + \frac{\partial^2 C}{\partial \eta^2}\right] \qquad (24)$$

where $\tau = t/t_{sc}$, $\alpha = x/a$, $\eta = y/\delta$, $C = c/c_{ref}$, and the expressions for u and v given by Eq. (13 b) and (17 b) have been used. In this flow, provided $Re_d >> 1$ we expect $(\delta/a)^2 << 1$. Also, because of the symmetry in the distribution of C about the stagnation point, a Taylor series expansion of this quantity in terms of α yields $C = C_0(\eta) + \alpha^2 C_2(\eta) + \ldots$, and we expect $\alpha\partial C/\partial\alpha \sim \alpha^2 << 1$. Thus, in the region around the stagnation point Eq. (24) simplifies to

$$\frac{\partial C}{\partial \tau} - \left[f + \frac{\phi}{2}\mathrm{Re}_d^{1/2}\right]\frac{\partial C}{\partial \eta} = \frac{1}{Sc}\frac{\partial^2 C}{\partial \eta^2}. \qquad (25)$$

Because f is known from the flow field analysis, this is a linear differential equation for C that is easily solved via a finite difference numerical procedure subject to appropriate boundary conditions, discussed below. The solution of Eq. (25) depends on Re_d, ϕ and the Schmidt number S_C ($\equiv v_f/D$), where D is the species diffusion coefficient in the fluid medium. Note that the dependence on the Reynolds number and the leakage parameter arises explicitly, through the $\phi\ Re_d^{1/2}$ term, and implicitly, through the dependence of the quantity f on these two parameters (see Eq. (21)). While there is no explicit dependence on a mass Peclet number ($Pe = Re_d\ Sc$) this can be made to appear, although in a rather artificial and awkward manner.

It should be noted that despite its apparent overly-simplified form, Eq. (25) is fully capable of accounting for the reduction in concentration due to the transport of odor species parallel to the antenna surface (in the direction parallel to x). This is a subtle point. A strictly one-dimensional convective diffusion equation, by default, is conservative in nature and conserves the total amount of odorant species. The only loss would come from the diffusion at the boundaries. This would be too restrictive a model of reality. In fact, the physical problem represented by such a one-dimensional model is one where an infinitely wide stream of air is depositing an infinitely wide filament of odorant on an infinitely-wide surface. The strength of the model captured in Eq. (25) is that the underlying convection is due to a true two-dimensional flow field (LHF). While a mathematically-rigorous analysis has demonstrated that the convective terms in the y direction can be neglected along the centerline, the velocity component *along* the centerline is not incompressible. Physically, there is loss of flow as it moves out laterally. The mathematical loss of flow is just what is needed to result in the correct behavior of the concentration field along the centerline.

2.3.1 Concentration field boundary conditions

The boundary conditions needed to solve the species concentration equation are obtained as follows. For the stagnation point flow mass transfer between an approaching stream and a surface, both at fixed and different concentrations, the ratio of the concentration boundary layer thickness to

the velocity boundary layer thickness varies according to $\delta_c/\delta \propto Sc^n$ with $n \approx -0.4$ (White 1974). Thus for the air flows of interest here $\delta_c < \delta$, meaning that the concentration boundary layer is embedded in the velocity boundary layer. Therefore, far away from the cylinder such that $\eta \to \infty$ meaning a value of y safely outside both the velocity and concentration boundary layers) the species concentration, in dimensional form, is prescribed as a time-dependent pulse of the form

$$c(y \to \infty, t) = c_\infty(t) = \frac{1}{2}\left(1 - \cos\frac{2\pi t}{T_P}\right), \quad 0 \le t \le T_P \tag{26}$$

with $c_\infty(t) = 0$ for all other values of time.

The boundary condition at the arc-shaped section of the cylinder surface corresponding to an antenna requires more careful consideration. In earlier studies for both animals in air and water (Adam and Delbrück 1968, Murray 1977, Stacey et al. 2002) the value of odorant species concentration in the fluid phase immediately adjacent to a sensillum surface modeled as a cylindrical surface has been taken as zero. This is equivalent to assuming instantaneous depletion of species at the sensillum surface which, in turn, presupposes infinitely fast diffusion and/or chemical reaction at that surface. Given the evidence in the literature for moth antennae (Kanaujia and Kaissling 1985), neither of these two conditions seems reasonable to us and we adopt a different, two-level mass-conservation approach. In this approach, for fluid mechanics purposes, we consider the sensilla to be small relative to the branches of the antenna so that, effectively, the antenna branches and their sensilla, together, form a single, continuous, porous mesh corresponding to the flow-exposed, arc-shaped surface in Fig. 1. At the same time, however, through a separate step, we retain the effects of the individual sensilla on the mass balance of chemical species.

This antenna model allows us to account for: a) the net convection and diffusion of species around and through the antenna; b) the adsorption and desorption of species on the sensilla; c) and, the diffusion of species into the sensilla. Diffusion of chemical species takes place through specialized pore-tubule units on the sensilla surfaces. Any amount of species not depleted from the flow passing through a leaky antenna emerges behind it. In this regard, to effect mathematical closure of the equations being solved, the flow emerging behind a leaky antenna is approximated as being purely convective (no diffusive component). The physical interpretation of this is that the concentration profile emerging behind the antenna is transported as an unchanged entity by the outflowing carrier fluid. Such an assumption is used widely in the computational modeling of elliptic/hyperbolic partial differential equations of which Eq. (25) is an example.

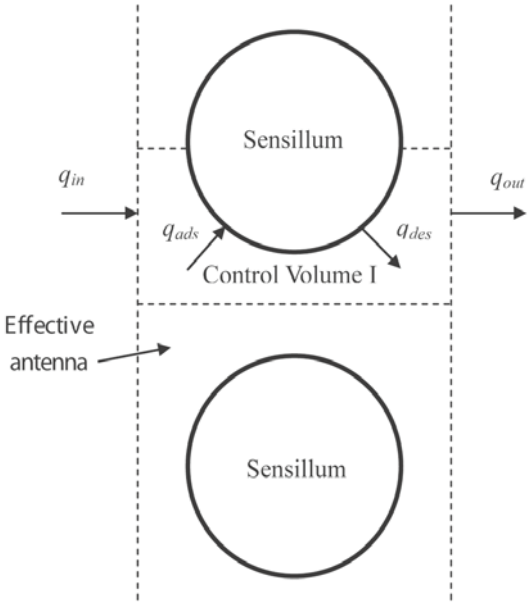


Fig. 3 Idealized schematic of a portion of the antenna (region of space contained between vertical dotted lines) and associated mass fluxes of odor species for Control Volume I of the analysis

With reference to Fig. 3, at the level of the antenna, a species mass balance performed on the elementary Control Volume I gives

$$q_{in} = q_{out} + (q_{ads} - q_{des})\frac{A_{ST}}{A_A}, \qquad (27)$$

where A_{ST} is the sum total of the area of the sensilla on the antenna, A_A is the area (including any open inter-branch and inter-sensilla spaces) of the antenna projected normal to the approaching flow. Two of the q terms represent species mass fluxes in and out of the antenna, and two are due to adsorption and desorption from the sensilla. The mass fluxes are given by

$$q_{in} = v_O c_O(t) + D\frac{\partial c_O(t)}{\partial y}, \qquad (28)$$

$$q_{out} = v_O c_I(t), \qquad (29)$$

$$q_{ads} = k_1\left(c_{S\max} - c_S(t)\right)c_I(t), \qquad (30)$$

$$q_{des} = k_2\, c_S(t), \qquad (31)$$

where: v_o is the prescribed fluid velocity passing through the antenna; $c_o(t)$ is the time varying concentration (mass/volume) of species in the approaching fluid, immediately adjacent to the antenna; $c_I(t)$ is the concentration (mass/volume) of species inside the control volume; c_S(t) is the time varying concentration (mass/area) of species on a sensillum surface; c_{Smax} is the maximum allowable concentration (mass/area) of species on a sensillum surface; and k_1 and k_2 are physical-chemical constants that must be determined experimentally, as must be c_{Smax}. In Eq. (30) and (31) we have adopted a Langmuir isotherm framework to model the adsorptive and desorptive fluxes. In this framework adsorption is proportional to the surface available for receiving the chemical species [$\propto (c_{Smax} - c_S$(t)] and to the chemical potential driving the species towards the surface [$\propto c_I(t)$]. Desorption is proportional to the surface coverage [$\propto c_S$(t)].

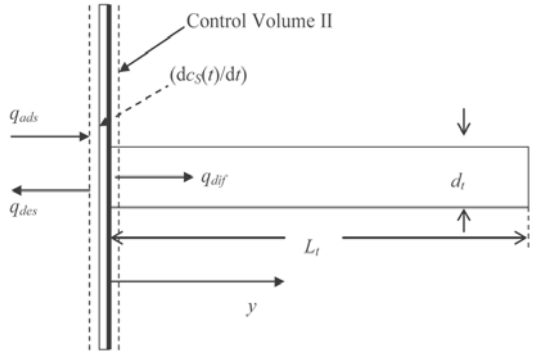


Fig. 4 Schematic of an idealized pore-tubule unit of diameter d_t and lenth L_t also showing the associated mass fluxes of odor species for Control Volume II of the analysis. The pore (left hand side of figure) sits on the surface of the moth sensillum. Odorant adsorption and desorption occur on the sensilla surface but only that fraction of species striking the pore diffuses to the end of the tubule

With reference to Fig. 4, at the level of the pore-tubule site for a moth through which diffusion takes place into a sensillum, a species mass balance performed on the elementary surface Control Volume II gives

$$\frac{dc_S(t)}{dt} = (q_{ads} - q_{des}) - q_{dif}\frac{A_{PT}}{A_S} \qquad (32)$$

where: A_{PT} is the sum total of the area of the sites (pores in the case of moths) on a single sensillum through which species diffusion takes place; A_S is the area of a single sensillum; and q_{dif} is the diffusive mass flux into the sensillum at the sites.

The quantity q_{dif} can be determined by solving the one-dimensional species diffusion equation in the direction normal to the sensillum surface subject to appropriate boundary conditions. In dimensional form, this equation and its boundary conditions are

$$\frac{\partial c}{\partial t} = D_t \frac{\partial^2 c}{\partial y^2} \tag{33}$$

with

$$t = 0, \qquad 0 \le y \le L_t,: \qquad c(y, 0) = 0 \tag{34}$$

$$y = 0, \qquad t \ge 0: \qquad c(0, t) = k_3\, c_S(t). \tag{35}$$

For the case of fast chemistry (diffusion-limited case) at the end of a pore-tubule system,

$$y = L_t, \qquad t \ge 0: \qquad c(L_t, t) = 0. \tag{36}$$

For the case of slow chemistry (kinetics-limited case),

$$y = L_t, \qquad t \ge 0: \qquad c(L_t, t)/\partial\, \mathrm{y} = 0. \tag{37}$$

In the above expressions, L_t is the length of the pore-tubule unit embedded in the moth sensillum. The quantity k_3 relates the species concentration (mass/volume) at $y = 0^+$, meaning inside the sensillum surface and immediately adjacent to it, to the species concentration (mass/area) on the sensillum surface. The quantity D_t represents the diffusion coefficient of the species in the tubule medium. Both k_3 and D_t must be determined experimentally. In this approach, it is tacitly assumed that the concentration of species on the sensillum surface, $c_S(t)$, remains uniformly distributed on the surface for all time. Also, if the rate of species diffusion into the sensillum surface is small relative to the accumulation of species at the surface, the value of $c_S(t)$ will not be significantly affected by the former process. This appears to be the case for moths (Kanaujia and Kaissling 1985) for which, in addition, $A_{PT}/A_S \cong 10^{-3}$ in Eq. (32), allowing the solution of Eq. (33) to be safely decoupled from that of Eq. (25).

2.4 Fluid properties, physical-chemical constants, and flow conditions

The model presented above contains a number of physical-chemical constants, area ratios, pulse duration, and maximum concentrations that must be specified in order to perform quantitative numerical calculations. In the case of moths, especially *Antheraea polyphemus*, there is enough related information available (Boeckh et al. 1965, Kanaujia and Kaissling 1985, Kaissling 1985, 1997) from which to obtain order of magnitude estimates for all quantities except k_3. For k_3 we arbitrarily set this quantity to the value unity. Fortunately, because the process of species diffusion through a sensillum can be decoupled from the process of species deposition on it, the exact value of k_3 is not indispensable to the calculations. When available, values obtained for k_3 can be used to derive more accurate values of species concentration in the tubule of a moth sensillum from the method provided here. In this regard, it is important to note that any uncertainty in k_3 affects the magnitude, but not the time-variation, of species concentration in the tubule or lymph.

Table 1 lists all the parameters required and conditions explored in this study. Although values for A_{ST}/A_A are provided in the table, because the solutions of Eq. (25) and (33) are decoupled this parameter is not actually used in the calculations. The properties for air correspond to a temperature of 27 °C.

Table 1 Flow conditions, values of the fluid properties, physical-chemical constants, maximum volume and surface concentrations, time pulse, and area ratios used in the analytical calculations. Many of these parameters represent best estimates derived indirectly from various sources of information available in the literature (Boeckh et al. 1965, Kanaujia and Kaissling 1985, Kaissling 1985, 1997) and are expected to provide correct typical qualitative trends. Fluid properties are from Incropera and De Witt (1990)

Parameter	Value (air at 27 °C)
ρ_f(fluid density; kg m^{-3})	1.16
ρ_f(fluid dynamic viscosity; kg m^{-1} s^{-1}	1.85×10^{-5}
v_f(= μ_f/ρ_f, fluid kinematic viscosity; m^2 s^{-1})	1.59×10^{-5}
d (effective antenna/antennule diameter; m)	6×10^{-3}
V_∞ (approaching or relative fluid speed; m/s)	0.5, 1
ϕ (= v_o/v_∞, effective antenna/antennule permeability)	0, 0.05, 0.1
Re_d (= $d\ v_\infty/v_f$, effective antenna/antennule Reynolds number)	189, 377
δ/d ($\cong Re_d^{-1/2}$, effective antenna/antennule dimensionless velocity boundary layer thickness)	0.05–0.07
D (odor diffusion coefficient in fluid; m^2/s)	2.5×10^{-6} (bombykol) 2.5×10^{-5} (other odors)
D_t (odor diffusion coefficient in tubule or lymph; m^2/s)	10^{-10}
Sc (= v_f/D, Schmidt number)	6.4 (bombykol) 0.64 (other odors)
δ_c/δ ($\cong Sc^{-0.4}$, effective antenna/antennule dimensionless concentration boundary layer thickness; estimate assumes stagnation point flow mass transfer)	0.48 (bombykol) 1.2 (other odors)
$c_{\infty max}$ (maximum concentration of odor pulse in fluid; gr/m^3)	10–8
c_{Smax} (maximum concentration of odor on sensillum surface; gr/m^2)	10^{-4}
L_t (length of tubule in moth sensillum or penetration depth in crayfish sensillum; m)	5×10^{-7}
A_{ST}/A_A (ratio of total sensillum area to effective antenna/antennule area normal to flow)	1
A_{PT}/A_S (ratio of total pore area on a moth sensillum to area of the sensillum; in the case of an aesthetasc there are no pores but entire sensillum surface participates in diffusion)	10^{-3}
T_P (odor pulse time; s)	5×10^{-3}
k_1 (sensillum adsorption constant; m^3 gr^{-1} s^{-1})	10^7
k_2 (sensillum desorption constant; s^{-1})	2.5×10^2
k_3 (outer-to-inner sensillum surface concentration equilibrium constant; m^{-1})	1 (assumed)

3. Numerical solution procedure

The governing equations are solved numerically. The equation for the leaky Hiemenz flow (Eq. (21)) is solved by discretizing the domain $\eta \equiv [0, \delta]$ into equally spaced intervals. The function $f(\eta)$ is evaluated at the node points. Given the third-order of the equation, and that there are two boundary conditions on $\eta = 0$ (antenna surface) and one on $\eta = \delta$ (free-stream), the boundary-value problem is solved iteratively by employing a shooting methodology. Shooting means that the boundary value problem is solved by imposing the two desired surface boundary conditions, plus a third, arbitrary condition on $f''(0)$. This formulation is then integrated via a fourth-order classical Runge-Kutta scheme in η until the free-stream is reached. The error in the desired free-stream boundary condition is then used as feedback to modify the arbitrary third boundary condition on the surface. This procedure is repeated until the free-stream boundary condition is satisfied to within a prescribed tolerance.

Once the Hiemenz function is determined, then the linear convective-diffusion equation for the concentration in the gas phase (Eq. (25)) is solved on the same grid used for the Hiemenz equation. The difference is that the species concentration is evaluated in the middle of each spatial interval (cell center), while the convective and diffusive fluxes are evaluated on the node points (cell faces). This system is discretized as a true initial-boundary-value problem. The discretization in time is via a fourth-order classical Runge-Kutta scheme. The discretization in space is second-order centered difference for the diffusive flux, and third-order upwind for the convective flux. The boundary condition at the free-stream is imposed as a flux boundary condition, i. e. the convective flux into the domain is taken to be $v_\infty\, c_\infty(t)$. The boundary conditions on the antenna surface involve the adsorptive and desorptive fluxes, as well as the surface concentration of the odor species on the antenna. The latter is governed by an evolution equation that is also solved via a fourth-order classical Runge-Kutta scheme.

In this work, the effect of the diffusion of odor species into the sensillum on the surface concentration of the odor species on the antenna is deemed to be negligible. Due to this decoupling, and due to the purely diffusive nature of the evolution of concentration along the tubule of a sensillum, the latter diffusion equation is solved accurately and efficiently using a standard Crank-Nicholson semi-implicit integration in time. This scheme allows taking time steps comparable to those employed in the flow domain. In this way, all concentration equations are then stepped in time together. An explicit scheme such as the Runge-Kutta would have resulted in prohibitively small time steps for the study of the diffusion process and would have lengthened the computations unreasonably.

4. Results and discussion

Numerical results have been calculated for the conditions listed in Table 1. A selection of the results obtained for moth antenna in air is presented and discussed in this section. For all cases a single odor pulse of duration $T_P = 5 \times 10^{-3}$ s arrives at the antenna. Two values of Re_d, two of S_c, and three of the leakage parameter, ϕ have been set. The two values of *Sc* arise because of the two different diffusion coefficients, *D*, investigated. The species diffusion coefficient is used to characterize diffusion through the moth sen-

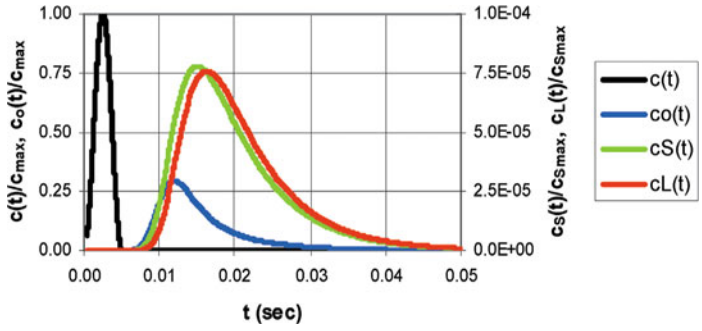


Fig. 5 Odorant species concentration as a function of time for a moth antenna in a flow with Re_d = 189, Sc = 6.4 and ϕ = 0.05. In this and all subsequent plots the following definitions apply: $c(t)$: fluid phase odor pulse concentration; $co(t)$: fluid phase odorant concentration immediately adjacent to the antenna surface; $c_S(t)$ surface concentration on the antenna; c_L(t): concentration at the end of the pore/tubule in the sensillum; c_{max}: maximum concentration in the odor pulse; c_{Smax}: maximum surface concentration on the antenna

sillum tubule yielding a characteristic diffusion time, $L_t^2/D_t = 2.5 \times 10^{-3}$ s. The maximum allowable surface concentration of odor corresponds to a single layer of molecules on the entire sensillum surface. The ratio of total sensillum area to the antenna area projected normal to the flow (including any open inter-branch and inter-sensilla spaces) is close to unity. In the moth about 10^{-3} of the sensillum surface consists of pores connected to tubules through which the odor species diffuses into the sensillum lymph.

Calculations have been performed assuming either very fast or very slow species depletion chemistry at the end of the pore tubule. In the case of fast chemistry relative to diffusion at position $y = L_t$ in the sensillum, the appropriate boundary condition to use is given by Eq. (36) and one obtains the time variation of species flux at L_t. In the case of slow chemistry, the appropriate boundary condition is given by Eq. (37) and the time variation of species concentration at L_t is obtained instead. The results provided here correspond to the slow chemistry assumption which best characterizes the problem of interest.

Typical results are shown in Fig. 5 to 8. The nominal case, shown in Fig. 5, corresponds to a flow with Re_d = 189, Sc = 6.4 and ϕ = 0.05. For this case, the initial odor concentration profile peaks at t = 2.5 ms while the air phase concentration immediately adjacent to the antenna surface peaks at about t = 12 ms and with a maximum value that is almost four times smaller than that of the pulse maximum. The reduction in concentration at this point is due to the transport of odorant species parallel to and away from the effective antenna surface as opposed to through it (see footnote 2). The antenna surface concentration grows simultaneously with that of the air phase adjacent to it and peaks at about t = 15 ms. Because of the smallness of the characteristic diffusion time of the odor species in the tubule (~ 2.5 ms), the concentration distribution at the end of the tubule tracks that on the antenna surface very closely. Notwithstanding, a time lag of 2–3 ms is observed between the respective peaks of these two curves and at about t = 17 ms they cross. This is because the surface is desorbed of species at a faster rate than diffusion can take place, now in the opposite direction, from the end of the tubule towards the sensillum surface.

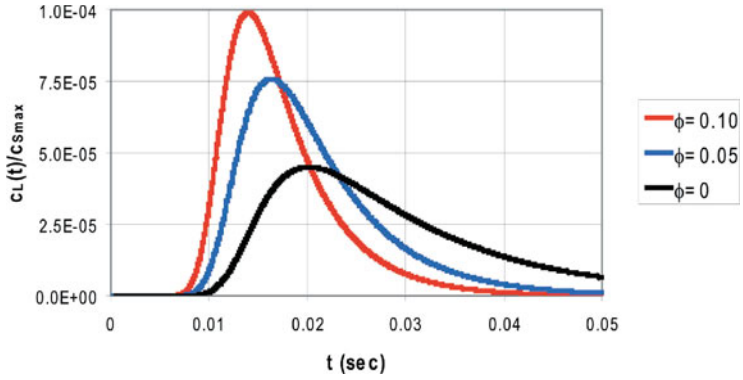


Fig. 6 Odorant species concentration as a function of time at the end L_t of a pore/tubule on the sensillum of a moth antenna in a flow with Re_d = 189 and Sc = 6.4 for ϕ = 0, 0.05 and 0.10. Same definitions as in Fig. 5 apply

By t = 30 ms the air phase species concentration adjacent to the antenna has decreased to very small values relative to its maximum but the surface concentration is still relatively large, correctly reflecting the expected slowness of the desorption process on the antenna (Kanaujia and Kaissling 1985). By t = 50 ms, both the surface and tubule-end species concentrations have dropped to very small values relative to their maxima.

The effect of antenna leakage is shown in Fig. 6 which provides plots of species concentration as a function of time at the end of the tubule on the sensillum of a moth antenna for the nominal flow (Re_d = 189, Sc = 6.4) with ϕ = 0, 0.05 and 0.10. From these results it is clear that increasing antenna permeability to flow results in shorter times for species concentrations to peak at larger values at the end of the tubule.

The effects of Sc and Re_d numbers are shown in Fig. 7 and 8. The consequence is that, because increasing Re_d and Sc respectively thin out their associated velocity (δ) and concentration (δ_c) boundary layers, small values of Sc (large D) and large values of Re_d (large v_∞) both favor the convective/diffusive transport of odorant species to the antenna surface and, ultimately, by diffusion, to the end of a tubule.

It is interesting to note that only a small fraction of the surface of a moth sensillum is covered in pore-tubule units capable of detecting an odor species like pheromone. However, the number of sensilla on an antenna is very large and, from Table 1, the area fraction available for odorant detection on an antenna is $A_{PT}/A_S \times A_{ST}/A_A \cong 10^{-3}$.

Because the Re_d and Sc numbers are functions of fluid properties, it is worth considering their theoretical variations with temperature. In air at constant pressure, we expect $\rho \propto T^{-1}$, $\mu \propto T^{1/2}$ and $D \propto T^{3/2}$. From these relations it follows that in air

$$\frac{\mathrm{Re}_d|_{T_1}}{\mathrm{Re}_d|_{T_2}} \cong \left(\frac{T_2}{T_1}\right)^{3/2}, \quad \frac{Sc|_{T_1}}{Sc|_{T_2}} \cong 1. \qquad (38)$$

From these ratios we see that the value of Sc in air is essentially independent of temperature. Taking T_1 = 283.2 °K (10 °C) and T_2 = 293.2 °K (20 °C) for illustration purposes, we find that $(\mathrm{Re}_d|_{10°}/\mathrm{Re}_d|_{20°})_{air} \cong 1.05$, $(\mathrm{Re}_d|_{10°}/\mathrm{Re}_d|_{20°})_{water} \cong 0.96$ and $(Sc|_{10°}/Sc|_{20°})_{water} \cong 1.14$. Such weak temperature dependencies are not expected to significantly influence the odorant species detection process.

The present study has been limited to single odor filament pulses of the half-cosine wave form given by Eq. (26). The question arises what would be the physical response of a sensillum exposed to a series of odorant pulses of randomly variable amplitude and period. This problem can be dealt with numerically but is computationally intensive. Instead, for guidance we look to a simi-

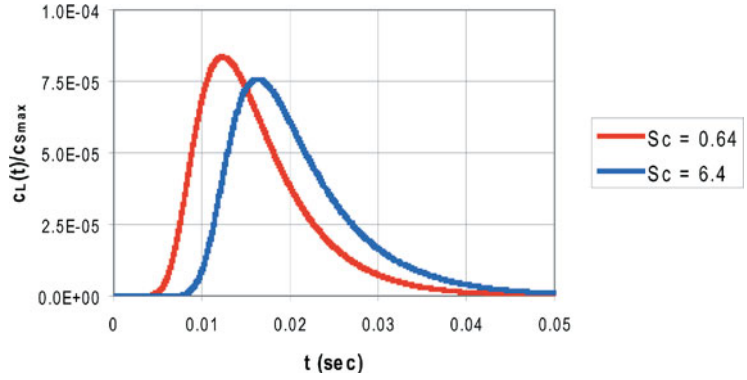


Fig. 7 Odorant species concentration as a function of time at the end L_t of a pore/tubule on the sensillum of a moth antenna in a flow with Re_d = 189 and ϕ = 0.05 for Sc = 0.64 and 6.4. Same definitions as in Fig. 5 apply

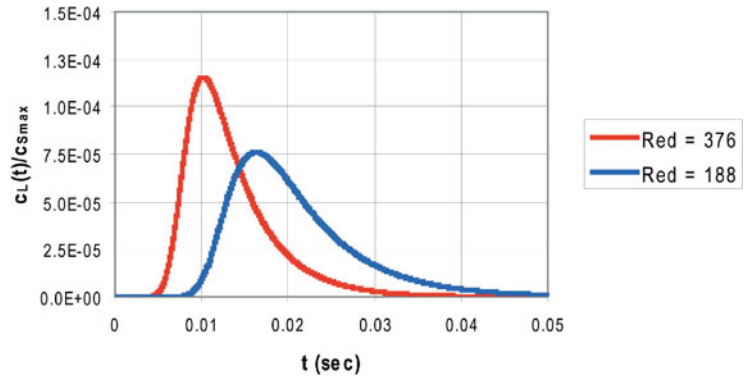


Fig. 8 Odorant species concentration as a function of time at the end L_t of a pore/tubule on the sensillum of a moth antenna in a flow with Sc = 6.4 and ϕ = 0.05 for Re_d = 189 and 377. Same definitions as in Fig. 5 apply

lar problem in the heat transfer literature, dealing with one-dimensional conduction into a semi-infinite solid with surface temperature periodic with time (Carslaw and Jaeger 1959; Eckert and Drake 1972). Applying the findings of this simpler problem to the pore-tubule unit of a moth sensillum shows that concentration oscillations at the sensillum surface will propagate into the sensillum with a wave velocity given by $\sqrt{2D_t\omega}$ and with decreasing amplitude given by $e^{-y\sqrt{\omega/2D_t}}$, where ω is the frequency in rad/s of the oscillation. Thus, while wave speed increases quadratically with oscillation frequency, the amplitude decreases exponentially, with the result that the amplitudes of oscillations with frequencies larger than about 20 Hz are significantly reduced. This physical finding is in agreement with neurosensory observations in the literature. For example, Kaissling (1997) states that both receptor cells and central neurons in male moths can resolve up to ten odor pulses per second.

We note that in our model the leakage parameter, $\phi = v_o/v_\infty$, is explicitly defined without needing to know the actual pressure drop required across the antenna to achieve the flow permeability specified. This does not pose a problem in those cases where the leakage is empirically known as a function of the antenna/sensilla geometry and the approaching fluid velocity (Vogel 1983). In a more elegant (but much more difficult) approach to the problem, a relationship is required that connects the antenna/sensilla geometry to the pressure drop across the antenna and the flow through it.

Present convection and diffusion transport time scales predicted for a moth antenna in a typical odor laden flow are in broad qualitative agreement with observations made in the literature. The conditions calculated correspond closely to the conditions

given in Fig. 3 in Kaissling (1997). In his study, Kaissling (1997) points out the existence of a 25 ms lag between the passage through an odor filament and the firing of action potentials in the sensillum receptor cells. Within 100 ms of passage through the filament, the moth initiates an anemotactic maneuver, and within 200 ms the receptor cells cease firing. The results in Fig. 5 show that within about 17 ms following initial contact with an odor filament pulse of 5 ms duration, the gas phase odorant concentration on the antenna surface has been significantly reduced ($c_0(t)/c_{max} < 0.12$) but that at the end wall of a tubule is just maximizing ($c_L(t)/c_{Smax} \cong 7.5 \times 10^{-5}$). While the absolute concentration threshold for odorant detection at the end of a tubule is not known, it is, presumably, quite small. Present results show that $c_L(t)/c_{Smax} < 10^{-6}$ for $t > 50$ ms.

While the fluid mechanics of the present problem dictate that the bulk of a filamentous odor plume should pass over a permeable antenna, that odor which leaks through the antenna generally contains more than enough molecules to stimulate the chemosensilla on the antenna.

The analysis presented here is readily extended to antennae/antennules of arthropods living in water. While the values of Re_d will not differ substantially between the two media, that for Sc in water is of order 1 000 compared to about 10 in air. Because $\delta_c/\delta \propto Sc^{-0.41}$ approximately, (White 1974) the larger values of Sc in water will result in much smaller values of δ_c/δ.

This study points to the need for improved or new experimental techniques for measuring (under carefully controlled but realistic conditions) the kind of data needed by predictive physical-mathematical models of the type presented here.

Acknowledgements

The authors are pleased to acknowledge stimulating and very helpful technical discussions with F. Barth, K.-E. Kaissling, and D. Mellon. In particular, we are most grateful for the considerable effort Professor Barth invested in helping us prepare this chapter to make it as appealing and accessible as possible to the biologist reader. A preliminary version of this work was presented as an invited lecture at the AFI-2002 Mini-Symposium on Advanced Fluid Information – Fusion of EFD and CFD, National Olympics Memorial Youth Center, Yoyogi Kamizono-cho, Tokyo, Japan, December 17, 2002. This chapter represents a significant revision of that lecture. JACH gratefully acknowledges funding received through NSF award CBET-0933034 in support of chemoreception research.

References

Adam G, Delbrück M (1968) Reduction in dimensionality in biological diffusion processes. In: Rich A, Davidson N (eds) Structural chemistry and molecular biology. Freeman and Co., San Francisco, pp 198–215

Baker TC (1986) Pheromone-modulated movements of flying moths. In: Payne TL, Birch MC, Kennedy CEJ (eds) Mechanisms in insect olfaction. Clarendon Press, pp 38–48

Barth FG, Schmid A (eds) (2001) Ecology of sensing. Springer, Berlin Heidelberg

Barth FG, Humphrey JAC, Secomb T (eds) (2002) Sensors and sensing in biology and engineering. Springer, Vienna New York

Batchelor GK 1967 An introduction to fluid dynamics. Cambridge University Press, Cambridge UK

Bell WJ, Cardé RT (eds) (1984) Chemical ecology of insects. Chapman and Hall Ltd. New York, NY

Berg HC (1983) Random walks in biology. Princeton University Press, Princeton NJ

Berg HC, Purcell EM (1977) Physics of chemoreception. Biophys J 20: 193–219

Bird RB, Stewart WE, Lightfoot EN (1960) Transport phenomena. John Wiley & Sons, Inc., New York

Bjostad LB, Gaston LK, Shorey HH (1980) Temporal pattern of sex pheromone release by female *Trichoplusia ni*. J Insect Physiol 26: 493–498

Boeckh J, Kaissling K-E, Schneider D (1965) Insect olfactory receptors. In: Cold Spring Harbor Sym-

posia on Quantitative Biology (Sensory receptors) Vol. XXX: 263–280

Bossert WH, Wilson EO (1963) The analysis of olfactory communication among animals. J Theor Biol 5: 443–469

Callahan PS (1975) Insect antennae with special reference to the mechanism of scent detection and the evolution of the sensilla. Int J Morphol & Embryol 4: 381–430

Cardé RT (1984) Chemo-orientation in flying insects. In: Bell WJ, Cardé RT (eds) Chemical ecology of insects. Chapman and Hall, London. pp 111–124

Cardé RT, Baker TC (1984) Sexual communication with pheromones. In: Bell WJ, Cardé RT (eds) Chemical ecology of insects. Chapman and Hall, London. pp 355–383

Cardé RT, Dindonis LL, Agar B, Foss J (1984) Apparency of pulsed and continuous pheromone to male gypsy moths. J Chem Ecol 10: 335–347

Carslaw HS, Jaeger JC (1959) Conduction of heat in solids, 2nd ed. Clarendon Press, Oxford

David CT (1986) Mechanisms of directional flight in wind. In: Payne TL, Birch MC, Kennedy CEJ (eds) Mechanisms in insect olfaction. Clarendon Press. Oxford UK. pp 38–48

Eckert ERG, Drake RM (1972) Analysis of heat and mass transfer. McGraw-Hill, New York NY

Elkington JS, Cardé RT (1984) Odor dispersion. In: Bell WJ, Cardé RT (eds) Chemical ecology of insects. Chapman and Hall, London. pp 73–81

Elkington JS, Cardé RT, Mason CJ (1984) Evaluation of time-average dispersion models for estimating pheromone concentration in a deciduous forest. J Chem Ecol 10: 1081–1108

Futrelle RP (1984) How molecules get to their detectors: the physics of diffusion of insect pheromones. Trends Neurosci 7: 116–120

Hansson BS (ed) (1999) Insect olfaction. Springer-Verlag, Berlin Heidelberg

Incropera FP, De Witt DP (1990) Fundamentals of heat and mass transfer (3rd ed.) John Wiley & Sons, Inc. New York NY

Kaisling K-E (1971) Insect olfaction. In: Beidler LM (ed) Handbook of sensory physiology 4. Springer-Verlag, Berlin Heidelberg New York. pp 351–431

Kaissling K-E (1985) R H Wright lectures on insect olfaction (Colbow K, ed.) Simon Fraser University, Burnaby, B. C., Canada, printed by Typographischer Betrieb W. Biering, H. Nürnberger, Munich, Germany

Kaissling K-E (1986) Chemo-electrical transduction in insect olfactory receptors. Ann Rev Neurosci 9: 121–145

Kaissling K-E (1997) Pheromone-controlled anemotaxis in moths. In: Lehrer M (ed) Orientation and communication in arthropods. Birkhäuser Verlag-Basel/Switzerland 343–374

Kaissling K-E (1998) Pheromone deactivation catalyzed by receptor molecules: a quantitative kinetic model. Chem Senses 23: 385–395

Kaissling K-E (2001) Olfactory perireceptor and receptor events in moths: a kinetic model. Chem Senses 26: 125–150

Kaissling K-E (2004) Physiology of pheromone reception in insects (an example of moths). Anir – Ανή (6: 73–91

Kaissling K-E (2009) Olfactory perireceptor and receptor events in moths: a kinetic model revised. J Comp Physiol A 195: 895–922

Kanaujia S, Kaissling K-E (1985) Interactions of pheromone with moth antennae: adsorption, desorption and transport. J Insect Physiol 31: 71–81

Karg G, Suckling M (1999) Applied aspects of insect olfaction. In: Hansson BS (ed) Insect olfaction. Springer-Verlag, Berlin Heidelberg. pp 352–377

Kennedy JS (1983) Zigzagging and casting as programmed responses to wind-borne odour: a review. Physiol Entomol 8: 109–120

Kennedy JS (1986) Some current issues in orientation to odor sources. In: Payne TL, Birch MC, Kennedy CEJ (eds) Mechanisms in insect olfaction. Clarendon Press. Oxford UK. pp 11–25

Krasnoff SB, Roelofs WL (1988) Sex pheromone released as an aerosol by the moth *Pyrrharctia isabella*. Nature 333: 263–265

Laue M, Steinbrecht RA (1997) Topochemistry of moth olfactory sensilla. Int J Insect Morphol & Embryol 26: 217–228

Loudon C, Koehl MAR (2000) Sniffing by a silkworm moth: wing fanning enhances air penetration through and pheromone interception by antennae. J Exp Biol 103: 2977–2990

Loudon C, Davis E (2005) Divergence of streamlines approaching a pectinate insect antenna: consequences for chemoreception. J Chem Ecol 13: 1–13

Mafra-Neto A, Cardé RT (1994) Fine-scale structure of pheromone plumes modulates upwind orientation of flying moths. Science 369: 142–144

Milne-Thomson LM (1968) Theoretical hydrodynamics. MacMillan Education Ltd. London UK

Morita H, Shiraishi A (1984) Chemoreception physiology. In: Kerkut GA, Gilbert LI, (eds) Comprehensive insect physiology biochemistry and pharmacology (Nervous system: sensory) 6. Pergamon Press. Oxford UK. 133–170

Murlis J, Jones CD (1981) Fine-scale structure of odour plumes in relation to insect orientation to distant pheromone and other attractant sources. Physiol Entomol 6: 71–86

Murlis J (1986) The structure of odor plumes. In: Payne TL, Birch MC, Kennedy CEJ (eds) Mechanisms in insect olfaction. Clarendon Press. Oxford UK. pp 27–38

Murlis J, Elkinton JS, Cardé RT (1992) Odor plumes and how insects use them. Annu Rev Entomol 37: 505–532

Murray JD (1977) Reduction of dimensionality in diffusion processes: antenna receptors of moths. In: Murray JD, Nonlinear differential equation models in biology, 83–127, Clarendon Press, Oxford UK

Mustaparta H (1984) Olfaction. In: Bell WJ, Cardé RT (eds) Chemical ecology of insects. Chapman and Hall Ltd. New York NY. pp 37–70

Panton RL (1996) Incompressible flow. John Wiley &Sons, Inc. New York NY

Payne TL, Birch MC, Kennedy CEJ (986) Mechanisms in insect olfaction. Clarendon Press. Oxford UK

Percy JE, Weatherston J (1974) Gland structure and pheromone production in insects. In: Birch MC (ed) Pheromones. North-Holland, Amsterdam. pp 11–34

Preiss R, Kramer E (1986) Pheromone-induced anemotaxis in simulated free flight. In: Payne TL, Birch MC, Kennedy CEJ (eds) Mechanisms in insect olfaction. Clarendon Press. Oxford UK. pp 69–79

Scheer J (2003) Night visions: The secret designs of moths. Prestel, New York NY

Schlichting H (1968) Boundary-layer theory. McGraw-Hill Book Company, New York NY

Schneider D (1964) Insect antennae. Annu Rev Entomol 9: 103–122

Schneider D (1969) Insect olfaction: deciphering systems for chemical messages. Science 163: 1031–1037

Schneider D (1986) Physiology of insect olfaction – reflections on the last 30 years. In: Payne TL, Birch MC, Kennedy CEJ (eds) Mechanisms in insect olfaction. Clarendon Press. Oxford UK. pp 38–48

Shuranova ZhP, Burmistrov IuM (1996) Orienting reaction in invertebrates. Neurosci Behav Physiol 26: 406–415

Stacey M, Mead KS, Koehl MAR (2002) Molecule capture by olfactory antennules: Mantis shrimp. J Math Biol 44: 1–30

Steinbrecht RA (1997) Pore structures in insect olfactory sensilla: a review of data and concepts. Int J Insect Morphol & Embryol 26: 229–245

Steinbrecht RA (1999) Olfactory receptors. In: Eguchi E, Tominaga Y (eds) Atlas of arthropod sensory receptors. Chapter V. Springer, Tokyo. pp 155–176

Vickers NJ, Baker TC (1992) Male *Heliothis virescens* maintain upwind flight in response to experimentally pulsed filaments of their sex pheromone (Lepidoptera: Noctuidae). J Insect Beh 5: 669–687

Vickers NJ, Baker TC (1994) Reiterative responses to single strands of odor promote sustained upwind flight and odor source location by moths. Proc Natl Acad Sci USA. 91: 5756–5760

Vogel S (1983) How much air passes through a moth antenna? J Insect Physiol 29: 597–602

White FM (1974) Viscous fluid flow. McGraw-Hill, Inc. New York NY

Willis MA, David CT, Murlis J, Cardé RT (1994) Effects of pheromone plume structure and visual stimuli on the pheromone-modulated upwind flight of male gypsy moths (*Lymantria dispar*) in a forest (Lepidoptera: Lymantriidae). J Insect Beh 7: 385–409

Wright RH (1958) The olfactory guidance of flying insects. Can Entomol 90: 81–89

Wright RH (1964) The science of smell. George Allen and Unwin Ltd. London UK

Wright RH (1982) The sense of smell, CRC Press. Boca Raton FL

Zacharuk RY (1985) Antennae and sensilla. In: Kerkut GA, Gilbert LI (eds) Comprehensive insect physiology, biochemistry and pharmacology. Pergamon Press, New York 6: 1–69

Mechanoreception

IV

Man-made versus biological in-air sonar systems

13

Herbert Peremans, Fons De Mey, Filips Schillebeeckx

Contents

List of abbreviations

ITD Inter-aural Time Differences, IID Inter-aural Intensity Differences, HRTF Head Related Transfer Function, CT computer tomography

Abstract

In this chapter we will argue that biologically inspired sonar systems, i. e. man-made systems that implement functional principles of their biological counterparts, are capable of significantly improving the performance of current in-air sonar systems. Instead of collecting large numbers of sonar range readings from multiple observation points and combining them into a reliable environment map we advocate the use of intelligent sonar sensors capable of extracting significantly more information from a single measurement. As an example of this bio-inspired approach we present a binaural sonar sensor capable of localizing reflectors in 3D-space using broadband spectral cues introduced by the emitter and receiver directional filters. Acoustic simulations indicate that duplicating the outer ears and combining them with an emitter that acts by directing emitted energy in the frontal direction should be sufficient to approximate the significant features of the directional properties of a real bat's sonar system. Localisation is performed by a template matching scheme whereby the spectrum of the received echo signal is compared with a set of stored spectral templates, one for every direction. This bio-inspired 3D localisation scheme was implemented on a robotic bat head and validated in a series of experiments. The results from these experiments show that both the monaural and the binaural spectral cues introduced by the emitter/receiver directional filters carry sufficient information to discriminate between different reflector locations in realistic noise conditions. The experiments further show that to track a moving spherical target with our robotic system spectral information from both receivers is required.

Herbert Peremans
University of Antwerp, Active Perception Lab
Prinsstraat 13, 2000 Antwerpen, Belgium
e-mail: herbert.peremans@ua.ac.be

1. Introduction

Because of their ease of use and their low cost man-made in-air sonar systems have been used extensively to provide a rudimentary environment perception capability to mobile robots that have to navigate autonomously.

An ideal navigation sensor generates accurate position information about all the objects in a large area at a fast measurement rate. In other words, it has a high information rate. The major drawback of conventional in-air sonar sensors such as the Polaroid range sensor (Maslin 1983), is precisely their low information rate. Indeed, a single measurement returns the distance to the closest object within the ensonified region only. Moreover, the measurement takes about 40 ms for a range of 6 m because of the slow speed of sound. Furthermore, the target position estimate based on the reading returned by the sensor is both highly uncertain and ambiguous as interpreted by the widely used sensor model for conventional in-air sonar. In this model (Fig. 1 a), it is assumed (i) that a reflector is located directly in front of the sensor (bearing angle θ) at a distance r derived from the time-of-flight of a probing ultrasound pulse and (ii) that the region in between the sensor and the target (the region labelled 'empty zone') is free of reflecting objects. Note that the width of the ensonified region, approximated by a cone with an opening angle of 30 degrees for the Polaroid transducer, gives rise to a large uncertainty on the estimated bearing angle. Indeed, the true target can be anywhere within the region labelled 'object zone'. In addition, this sensor model gives rise to many erroneous interpretations of the sensor readings by ignoring the complexity of the reflection process. Figure 1 b shows how the presence of an obliquely oriented extended reflector acting like a mirror can give rise to such an erroneous interpretation. In this example, the mirror image of the target is mistaken by the sensor model for the true target. Because of these limitations the use of in-air sonar has been limited mostly to low-level obstacle avoidance tasks in robotics.

Bats, on the other hand, rely upon their airborne sonar system while executing complex navigation and hunting behaviour (Griffin 1958). Clearly, they have developed strategies to overcome the low information rate of in-air sonar sensing. These strategies fall into two broad classes: selective exploration and information extraction maximization. In

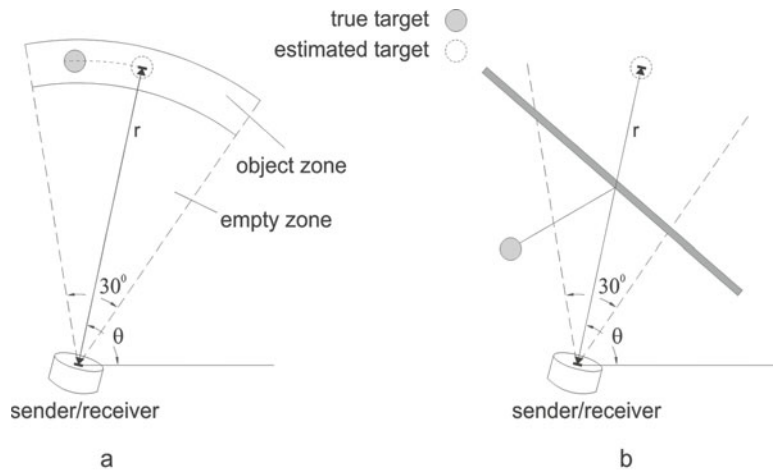


Fig. 1 **a** The estimated target position (r,θ) and its uncertainty (indicated by the label 'object zone'); **b** an erroneous measurement interpretation, based on the same sensor reading, mistaking the mirror image for the true target

this chapter, we will be concerned with the latter mainly.

As the speed of sound poses a physical constraint upon the measurement rate, given a certain spatial range of interest, bats have learned to adapt the configuration of their sonar to the task at hand. This allows them to actively explore their environment and renders possible more efficient extraction of information (Surlykke et al. 2009). In addition, bats have learned to extract much more detailed information from a single. Evidence for such extensive information extraction is provided by the various target feature maps (Suga 1990) found in the bat's auditory cortex: echo delay represents target range, echo amplitude+delay represent target size, echo amplitude spectrum represents target shape, binaural echo cues represent target azimuth+elevation, pinna induced monaural/binaural spectral cues represent target azimuth+elevation, constant Doppler shift component represents target relative velocity, periodic Doppler shift component represents target flutter, etc.

In this chapter we will argue that biologically inspired sonar systems, i. e. man-made systems that implement functional principles of their biological counterparts, are indeed capable of significantly improving the performance of man-made in-air sonar systems.

2. Man-made in-air sonar systems

To compensate for the low information rate of conventional man-made sonar systems, engineers have followed two approaches. In one approach, improved accuracy and reliability of the results is obtained by scanning the region of interest from different viewpoints with the basic range sensor. Hence, number of measurements (= time) is traded for quality of the results. With this approach, multiple sensors are often used in order to get more acceptable response times despite the low measurement rate of individual sensors. In the other approach, the information rate of a single intelligent sonar sensor is increased by extracting significantly more information from a single measurement.

2.1 Post-processing multiple range readings

The aim of the post-processing approach is to combine the unreliable sonar readings collected at multiple observation points and turn them into a reliable environment map using decision theoretic tools (Choset et al. 2005) It requires the specification of the following essential components.

(1) A model S of the environment to represent the acquired environment information. Note that this model includes the estimate of the robot's state, e. g. the mobile robot's position and orientation in the plane. The environment model can consist of geometric primitives such as lines, planes etc., whose uncertainty is modelled by interpreting their parameters as random variables with associated probability distributions as first proposed by Smith and Cheeseman (1986). Moravec and Elfes (1985) introduced the other frequently used environment model consisting of an occupancy/certainty grid. In this case, the map primitives are rectangular cells. The value corresponding to each cell denotes the degree of confidence in the cell being occupied by an object.

(2) A sensor model $P(M|S)$ defined as the conditional probability of the sonar reading M given the state of the environment S. The sensor model includes both

the effects of measurement uncertainty as well as the effects of ambiguous readings, i. e. identical sensor readings produced by different environment states (Fig. 1). Hence, denoting the different measurement scenarios consistent with the environment state by $C(S)$, the sensor model can be written as

$$P(M|S) = \sum_{\forall C(S)} P(M, C(S)|S) \quad (1)$$

with

$$P(M, C(S)|S) = P(M|S, C(S)) \cdot P(C(S)|S). \quad (2)$$

Note that the probability distribution $P(C(S)|S)$ quantifies the amount of ambiguity present in a particular measurement given the state of the environment S. $P(M|S,C(S))$, on the other hand, describes the uncertainty proper of the sonar measurement given the state of the environment and a particular measurement scenario, e. g. an object is present in the object zone determined by the beamwidth of the transducer together with the measured distance from the sensor and no object is present in the empty zone.

(3) An inferencing technique like Bayesian inference, to consistently combine the information extracted from the new measurement with the information collected so far (Thrun 1998).

The fundamental problem with the post-processing approach is that the sonar sensor model has to be simplified considerably to keep the procedure computationally tractable. Indeed, in practice it is not feasible to consider all measurement scenarios $C(S)$. As a result, faulty interpretations of ambiguous measurements will occur. Unless large numbers of sonar readings are collected so that the erroneous sensor-data interpretations get eventually erased, the results will still be unreliable.

2.2 Intelligent sonar sensor readings

The post-processing approach applied to conventional sonar sensors throws away a lot of information during the measurement process, only to attempt to reconstruct it afterwards at great expense. While this might be a sensible approach for fast sensors such as cameras or laser-range finders, it is not for sonar. Indeed, as argued before, the low propagation speed of sound makes the collection of large numbers of measurements very time consuming.

Starting with a more detailed look at the echo formation process, we describe a more intelligent bio-inspired in-air sonar system below.

2.2.1 Target feature encoding

Before discussing the extraction of information from received echo signals, we first look at how environmental features get encoded into particular features of those signals.

The measurement using an in-air sonar system consists of the following steps: a sound pulse is generated and radiated into the environment. It travels through the air and reflects from all objects within range. The reflected echoes are picked up by the receiver(s). Modelling the transformations occurring at the different stages by linear processes (Kuc and Siegel 1987; Peremans 1997) the received signal $r(t)$ is given by

$$r(t) = \sum_{i=1}^{N_{refl}} h_{rec}(t;\overline{\theta}_i) * h_{refl,i}(t) * h_{air}(t;r_i) * h_{emit}(t;\overline{\theta}_i) * s_{call}\left(t - \frac{2r_i}{v_{sound}}\right), \quad (3)$$

with $s_{call}(t)$ representing the generated signal, $*$ denoting convolution, and the $h(t)$'s denoting the impulse responses corresponding with the different filter operations, i. e. the emission and reception directivity $h_{emit}(t;\bar{\theta})$ and $h_{rec}(t;\bar{\theta})$ and the spherical spreading and absorption losses in air $h_{air}(t;r_i)$. The information to be extracted from this composite echo signal consists of the positions of the reflectors of interest, given by spherical coordinates $(r,\bar{\theta})i = (r,\theta_{az},\theta_{elev})$, as well as their identities, given by the impulse responses $h_{refl,i}(t)$.

In the remainder of this chapter we will be concerned with inverting the transformations applied during the echo formation process to extract reflector position information from the relevant features of the received echo signals.

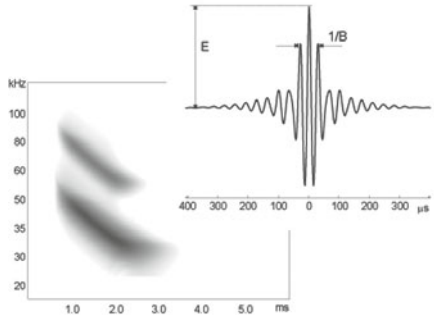


Fig. 2 The spectrogram (magnitude) of the emitted call; (inset) the output of the corresponding matched filter (E and B denoting respectively the energy and the bandwidth of the emitted call)

2.2.2 Extracting range information

Contrary to the classic range sensor we will attempt to localise all reflectors in the ensonified region. Hence, we extract the travel times of all the echoes and not just that of the first echo. In order to arrive at the minimum variance estimate of the arrival time of a single echo in the presence of additive white Gaussian noise, we have to process the received signal appropriately. From radar theory (Skolnik 2008) we know we have to use a so-called matched filter for this processing. The impulse response of a matched filter is given by the time inverse of the echo-signal. The arrival time estimate itself is given by the moment the matched filter output is maximal. The important features of the matched filter output are shown in Fig. 2. The echo signal shown is loosely modeled on the emitted call of *Eptesicus fuscus* (Surlykke and Moss 2000) in the terminal phase. It is a hyperbolic frequency modulated sinusoidal containing 2 harmonics with the fundamental component sweeping from 50kHz down to 25kHz in 2 ms. Note that the effective width of the matched filter response for a single echo is determined by the bandwidth and not by the duration of the emitted call.

Assuming the echo signal to be composed of multiple scaled and delayed versions of the emitted call and introducing a heuristic peak detection algorithm (Peremans et al. 1993) that can cope with multiple overlapping echoes, we can apply the same approach to realistic echo signals such as $r(t)$ specified above (see equation (3)).

As shown in Fig. 3 for a binaural sonar system consisting of one transmitter and two receivers, applying this approach to both echo signals results in two sets of echo arrival times.

The arrival times corresponding to the echoes from the same reflector picked up by the right and left receivers contain information about both the spatial range and the bearing of the reflectors in the horizontal plane. The range can be derived from the arrival times themselves. The bearing can be inferred by triangulation from the difference between the arrival times at both ears, i. e. the inter-aural time difference (ITD). A number of robotic sonar systems have been proposed that use the arrival times and ar-

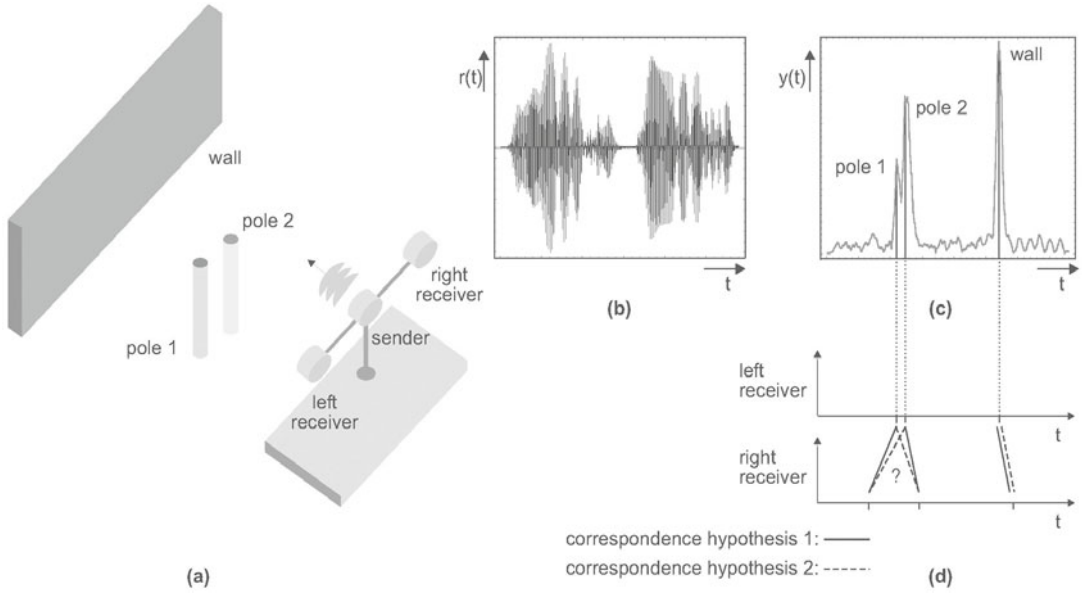


Fig. 3 **a** The environment and binaural sonar system; **b** the echo signal from the left receiver; **c** the output of the corresponding matched filter; **d** the sets of echo arrival times extracted from the matched filter outputs for the left and right receivers and the associated correspondence hypotheses

rival time differences of echoes extracted from multiple receivers to localise objects (Peremans et al. 1993; Kleeman 2004). When multiple arrival times are extracted from each receiver signal in these systems, a matching problem quite similar to the correspondence problem in stereo-vision needs to be solved. Indeed, the echoes in the right and left receiver signals belonging to the same reflector need to be paired (Peremans 1994) before triangulation can be used to determine the bearings of the individual reflectors. For example, in Fig. 3 d the third echo in the left receiver signal corresponds unambiguously with the third echo in the right receiver signal. However, both the first and the second echo in the left receiver can correspond with either the first or the second echo in the right receiver. This results in two correspondence hypotheses: hypothesis 1 pairs the first echo in the left receiver with the first echo in the right receiver and the second echo in the left receiver with the second echo in the right receiver whereas hypothesis 2 pairs the first echo in the left receiver with the second echo in the right receiver and the second echo in the left receiver with the first echo in the right receiver. In both hypotheses the third echo in the left receiver is paired with the third echo in the right receiver. Of the two correspondence hypotheses only one will result in correctly triangulated position estimates for the first and the second target. Since reliably finding the true correspondences becomes more complicated as the receivers are further apart it would seem to be advantageous to build small sonar systems. However, the angular accuracy of a position estimator based on ITD cues is inversely proportional to the distance between the receivers. Indeed, it has been suggested that the small size of the bat's head makes inter-aural intensity (IID) cues much more reliable than ITD cues for target localisation (Pollak 1988). Furthermore, when confronted with a true 3 D environment ITD cues from two receivers contain

only partial information about the reflector positions.

Hence, instead of introducing additional receivers and limit ourselves to processing echo arrival times, we take inspiration from biology again and turn our attention to a different set of position cues. In particular, to localise reflectors in 3 D while keeping with a binaural sonar system we propose to use the cues introduced by the emitter and receiver directional filters.

2.2.3 Extracting angular information

The direction dependent filtering performed by the diffraction and interference of the echo sound field around the receiver and the structures close to it is referred to as the Head Related Transfer Function (HRTF). A similar direction dependent filter is present during emission of the call as well. For bats, the complexity of the shapes involved (pinnae, head, shoulders, wings, etc.) excludes the analytic derivation of the relevant sound fields. Hence, until recently bat HRTF's (Firzlaff and Schuller 2003; Aytekin et al. 2004) could only be measured. However, during the last few years numerical tools have become available that allow to simulate the acoustic field around complex structures. Recently, we have developed a method (De Mey et al. 2008) that allows to predict the filters $h_{emit}(t;\bar{\theta})$ and $h_{rec}(t;\bar{\theta})$ with reasonable accuracy, given a particular receiver structure. The method starts with a semi-automatic extraction of the shape of a complete bat head from micro-CT data. Next, this detailed morphology data is used in combination with numerical tools to simulate the acoustic field.

Apart from the efficiency gain, this new approach allows to investigate what parts of the bat head contribute significantly to

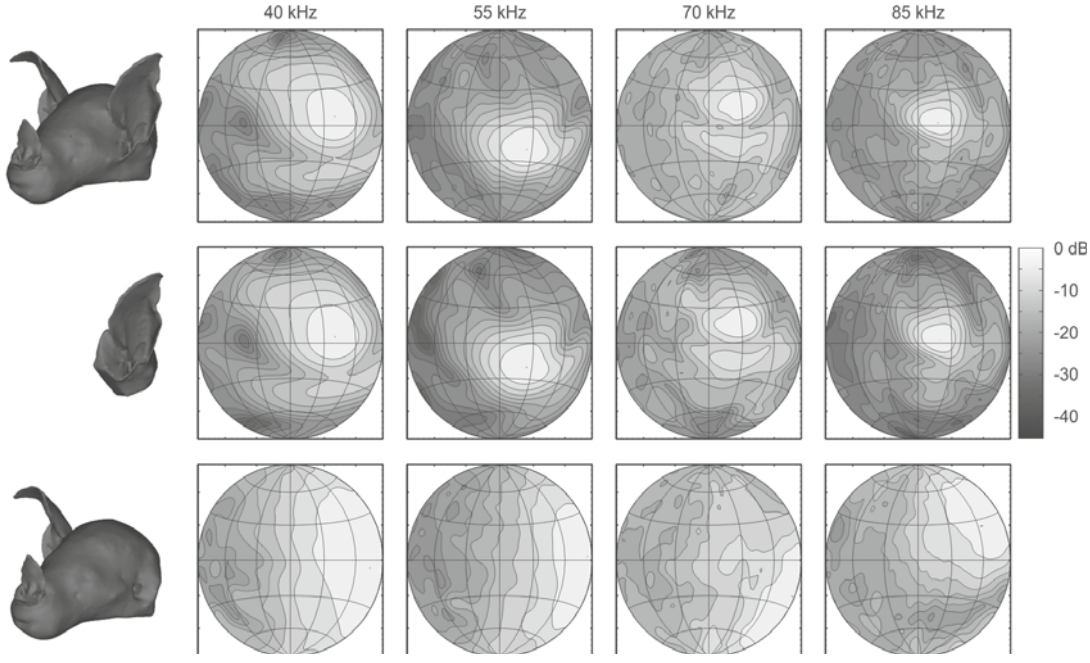


Fig. 4 The directional dependency (Lambert azimuthal equal-area projection) of the reception filters for different frequencies and different configurations: complete head, isolated pinna, head without pinna. The spatial sensitivity patterns are normalised such that 0 dB corresponds with the highest sensitivity for each frequency, the contours are spaced 3 dB apart (grid lines every 30 degrees)

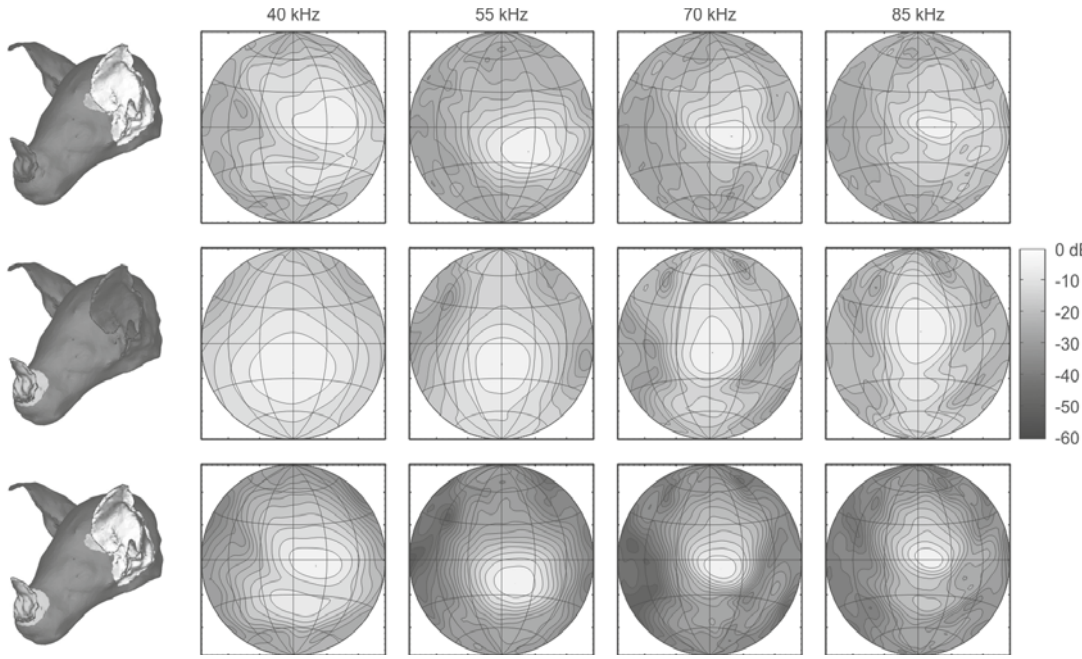


Fig. 5 The directional dependency (Lambert azimuthal equal-area projection) of the reception filter, the emission filter and the combined sender–receiver filter for different frequencies. The patterns are normalised such that 0 dB corresponds with the highest value for each frequency, the contours are spaced 3 dB apart (grid lines every 30 degrees)

the measured directional filters. In particular, from the results shown in Fig. 4 we conclude that the contribution of the pinna to the HRTF is much more important than that of the rest of the head for *Phyllostomus discolor.* Indeed, the important properties of the spatial sensitivity patterns of the complete head, i. e. both the size and position of the main lobe as well as the presence/absence of significant side lobes as a function of frequency, are well predicted by the spatial sensitivity patterns of the pinna considered in isolation. At the same time, the spatial sensitivity patterns of the complete head and the head with the pinna removed are only very weakly correlated.

As biosonar is an active sensing modality we need to extend the analysis to include sound emission as well. The middle row of Fig. 5 shows how in the case of *Phyllostomus discolor* the sound emitted through the two nostrils interacts with the noseleaf to create an elongated main beam in front of the bat. We conclude from Fig. 5 that combining the emission (middle row) and reception (top row) filters results in a more directional filter for the complete system (bottom row). Indeed, the sensitivity in the periphery of the combined sender-receiver system is significantly reduced with respect to that of the reception subsystem considered on its own. Nevertheless, the directional dependency of the combined filter as quantified by the size and position of the main lobe and most important side lobe has stayed quite close to that of the reception filter.

To extract the angular cues introduced by this combined sender-receiver directional filter in the received echo signals, we propose a simple template matching scheme (Reijniers and Peremans 2007). In this scheme the time-localised spectrum of the received

echo signal is compared with a set of stored spectral templates, one for every direction. The spectral template that has the smallest weighted distance, using the diagonal elements of the noise covariance matrix as the weights, with respect to the echo spectrum specifies the direction estimate. The analysis presented by Altes (1978) shows that this procedure results under general conditions in a minimum variance estimate of the direction in the presence of isotropic white Gaussian noise.

This scheme can be applied to the monaural as well as the binaural echo spectrum, i. e. the concatenation of the two monaural spectra. The monaural version has the advantage of not requiring the solution of the correspondence problem, i. e. finding the echoes in the right and the left receiver signals, that correspond with the same reflector. However, it lacks the capability of discriminating between the spectral features due to the reflector filter and those due to the emission and reception filters. Hence, the binaural version is more robust in the case of more pronounced reflector filtering.

3. Bio-inspired in-air sonar

The template matching scheme, which we introduced in the previous section, has been evaluated with an engineered sonar system. This section elaborates on the system, as well as on the obtained results.

3.1 Robotic bat head

The robotic bat head (Fig. 6) used in this research is based on the robotic head (Peremans and Reijniers 2005) developed as part of the CIRCE project (CIRCE 2005) that was

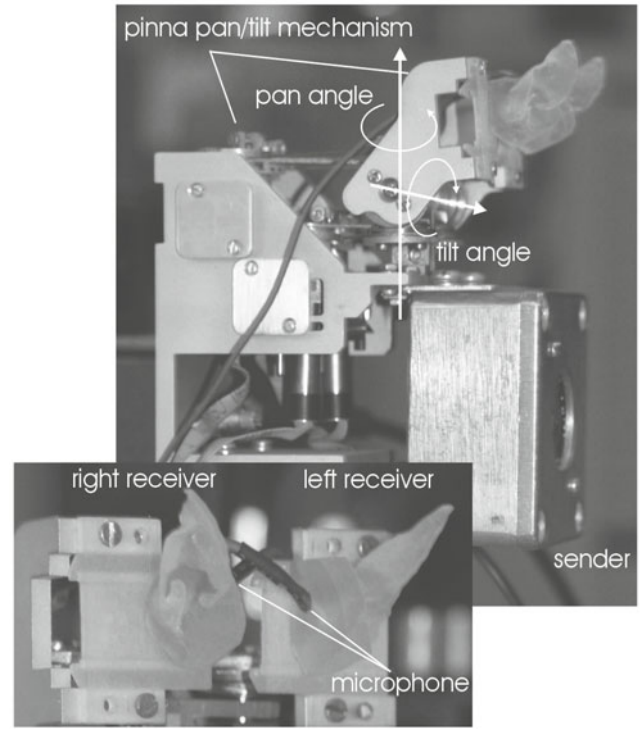


Fig. 6 The CIRCE head fitted with Polaroid transducer as sender and Knowles microphones mounted in plastic replicas of the *Phyllostomus discolor* pinnae as receivers

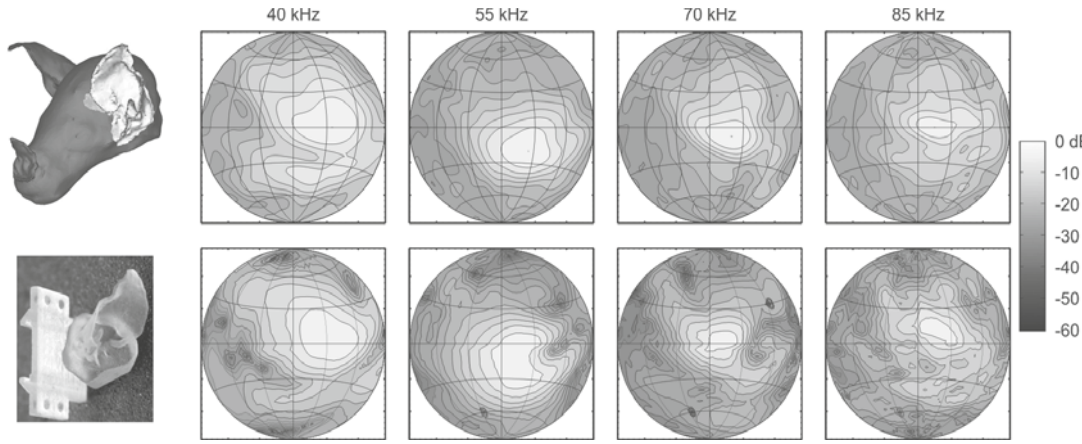


Fig. 7 Directional dependency of the reception filter of **(top)** the simulated *Phyllostomus discolor* HRTF; **(bottom)** the Knowles microphone inserted in the artificial *Phyllostomus discolor* pinna. The spatial sensitivity patterns are normalised such that 0 dB corresponds with the highest sensitivity for each frequency, the contours are spaced 3 dB apart (grid lines every 30 degrees)

concerned with reproducing, at a functional level, the echolocation system of bats.

Two microphones are mounted onto a micromechanical system that allows accurate independent control of both pan and tilt angle of the two receivers. Various pinna shapes can be mounted over each microphone. The cable driven mechanism results in a sensor work space of ±30 degrees in both the pan and the tilt direction. The micromechanical drive system is located behind the movable pinnae, reducing reflection and diffraction effects on the received signals.

For the experiments described below two microphones (Knowles FG-23 329) are inserted in the ear canals of the plastic pinnae mounted on the robotic bat head. The microphones' small active surface results in a spatial directivity that is almost omnidirectional for the frequency range of interest (30–95kHz). However, when inserted in the ear canal of the artificial *Phyllostomus discolor* pinnae, the microphones inherit the directional dependency of the true *Phyllostomus discolor* HRTF as can be seen from Fig. 7.

The transmitter is based on a Polaroid ultrasonic transducer driven by custom built electronics. Clearly, the directional filtering of this transducer (Peremans 1997) is different from that of the emission system of *Phyllostomus discolor.* However, from the simulation results shown in Fig. 5 we conclude that duplicating the outer ears and combining them with a directional emitter that acts by directing emitted energy in the frontal direction should be sufficient to approximate the significant features of the directional properties of the real bat's sonar system.

3.2 Experimental results

The bio-inspired 3D localisation scheme described above was implemented on the robotic bat head and validated in a series of experiments. We tested three different localisation strategies (monaural, selected monaural, and binaural) and show their localisation capabilities as a function of the position of the target in Fig. 8. The monaural strategy localizes echoes with information from one ear only, the right ear in this case. The strategy called *selected monaural*

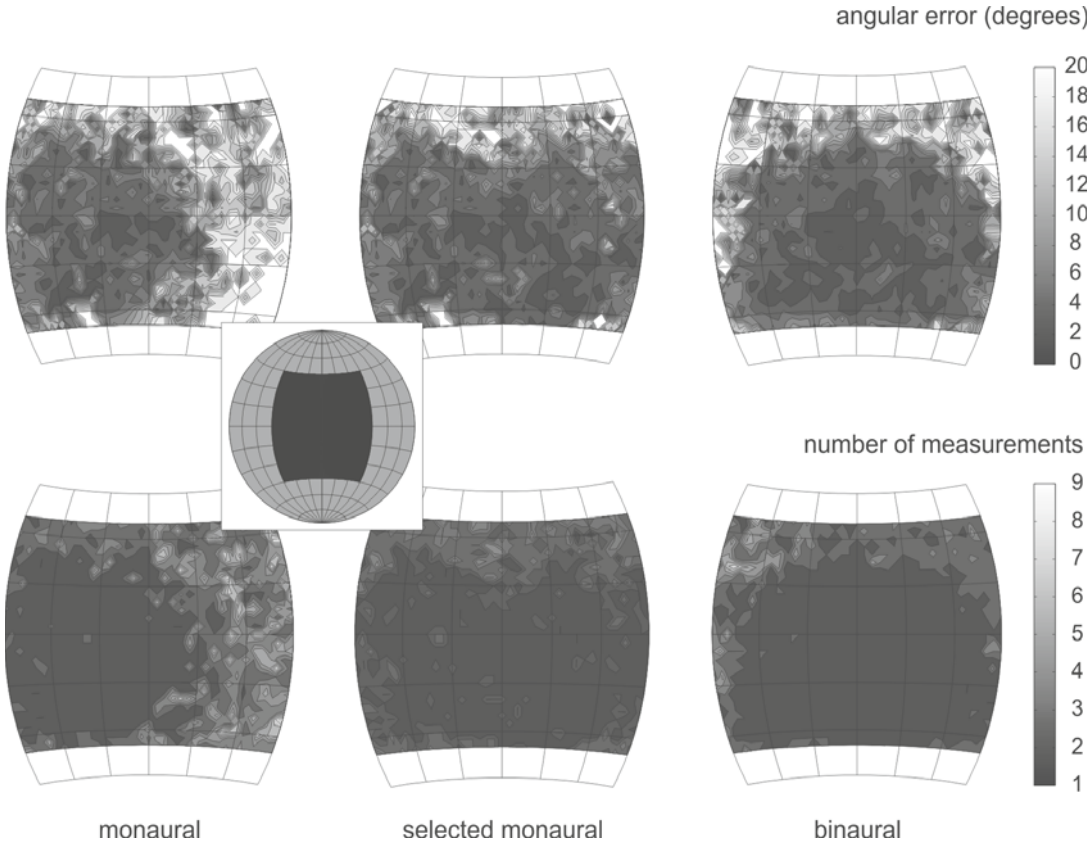


Fig. 8 Spatial distribution of angular estimate error in degrees (top); number of measurements to converge on target (bottom) making use of monaural information, selected monaural information (see text) and binaural information (grid lines every 15 degrees; inset shows which part of the frontal hemisphere is reported on)

first checks which ear receives the strongest echo and then applies the *monaural* strategy to the signal received by that ear. The strategy called *binaural* uses all information contained in the signals from both ears to perform true binaural localisation as explained in section 2.2.3. The target was a spherical reflector placed at approximately 1 m from the robotic bat head. The performance of each strategy was quantified by measuring the great-circle distance from the estimated target position to the true one (see Fig. 8, top row).

This error measure does not provide information about the possibility of sequential position estimation. Hence, we have also calculated (see Fig. 8, bottom row) how many sequential measurements it would take the system to arrive within a 3 degrees capture region around the true target position. This performance measure provides information about the stability and the rate of convergence of the robotic bat head tracking the target based on its angular position estimates. This performance measure seems more appropriate for a tracking task, as it takes into account that even if the first location estimate is quite far off, steering the robotic bat head to point in this faulty direction might relocate the target within a more favorable HRTF region. Subsequent measurements would then allow correct lo-

calisation. The maximum number of measurements in this experiment is set to 9.

Both performance measures show that monaural localisation of a single stationary target is possible, albeit with a reduced accuracy, as long as the target stays within the corresponding half of the region in front of the sonar system. However, it was found that this strategy was not capable of tracking a moving target. Indeed, the large errors it makes when the target is in the wrong frontal half space causes it to lose track of the moving target. Localisation in both halves of the frontal hemisphere ensonified by the emitter is possible by combining the readings from both receivers. Note that in this experimental setup finding the correspondence between the echoes from the right and the left receiver is trivial as there is only one echo. Both the *selected monaural* and the *binaural* strategy are capable of tracking a moving target in real-time on our robotic system. As can be seen from the results shown in Fig. 8 the overall accuracy of the position estimate, in particular in the region directly in front of the sonar system, is slightly higher in the *binaural* strategy. The *selected monaural* strategy has better performance for larger azimuth angles, i. e. a larger effective field of view. Note that performance for all three strategies deteriorates as the elevation angle gets larger. This is a consequence of the relative positioning of the emitter and receiver directivity beams. The combined emitter and receiver directivity of the sonar system is pointing slightly downward in this case.

Conclusion

We have argued that biologically inspired sonar systems are capable of significantly improving the performance of man-made in-air sonar systems. This can be achieved by copying some of the strategies used by bats to overcome the low information rate of in-air sonar sensing. In particular, we have shown that the extraction of binaural spectral information of the echo signals in combination with the use of a biologically inspired directional sonar system allows 3D localisation of multiple reflectors from a single measurement.

However, there is still a significant performance gap between biological sonar systems and man-made ones. In particular, despite the good results in the presence of multiple reflectors, the 3D localisation mechanisms proposed here break down when presented with too many densely spaced reflectors. In those circumstances the spectral properties of the echoes are no longer solely determined by the directional filtering of the sonar system but also by the target impulse response. The consequent additional spectral features tend to confuse the localisation mechanism. This often occurs when presented with environments containing natural reflectors such as trees. We are currently investigating to what extent the use of an intelligent focus-of-attention mechanism or the processing of trains of measurements instead of single measurements might remedy some of these remaining problems.

Acknowledgements

This work is supported by the EU, IST-FET Program, as part of the CILIA project and the BOF-UA Program, supporting interdisciplinary PhD-research.

References

Altes R (1978) Angle estimation and binaural processing in animal echolocation. J Acoust Soc Am 63: 155–173

Aytekin M, Grassi E, Sahota M. Moss CF (2004) The bat head-related transfer function reveals binaural cues for sound localization in azimuth and elevation. J Acoust Soc Am 116: 3594–3605

Choset H, Lynch KM, Hutchinson S, Kantor G, Burgard W, Kavraki LE, Thrun S (2005) Principles of robot motion. The MIT Press, Cambridge London

CIRCE (2005) URL: http://www.ua.ac.be/main.aspx?c=.APL&n=40656

De Mey F, Reijniers J, Peremans H, Otani H, Firzlaff U (2008) Simulated head related transfer function of the phyllostomid bat *Phyllostomus dis-*

color. J Acoust Soc Am 124: 2123–2132

Firzlaff U, Schuller G (2003) Spectral directionality of the external ear of the lesser spear-nosed bat, *Phyllostomus discolor.* Hear Res 181: 27–39

Griffin DR (1958) Listening in the dark: the acoustic orientation of bats and men. Yale University Press, New Haven

Kleeman L (2004) Advanced sonar with velocity compensation. Int J Rob Res 23(2): 111–126

Kuc R, Siegel M (1987) Physically based simulation model for acoustic sensor robot navigation. PAMI, 9(6): 766–778

Maslin GD (1983) A simple ultrasonic ranging system. In: 102nd Convention of AES, Cincinatti

Moravec H, Elfes A (1985) High resolution maps from wide angle sonar. In: Proc of the IEEE Int Conf on Robotics and Automation, St. Louis

Peremans H (1994) A maximum likelihood algorithm for solving the correspondence problem in tri-aural perception. In: Proc IEEE Int Conf on Multisensor Fusion and Integration for Intelligent Systems, Las Vegas

Peremans H (1997) Broad beamwidth ultrasonic transducers for tri-aural perception. J Acoust Soc Am 102(3): 1567–1572

Peremans H, Audenaert K, Van Campenhout JM (1993) A high- resolution sensor based on tri-aural perception. IEEE Trans Robotics and Automation 9(1): 36–48

Peremans H, Reijniers J (2005) The CIRCE head: a biomimetic sonar system. In: Artificial neural networks: Biological inspirations – ICANN 2005, LNCS 3696, Springer Verlag, Berlin

Pollak GD (1988) Time is traded for intensity in the bat's auditory system. Hear Res 36(2–3): 107–124

Reijniers J, Peremans H (2007) Biomimetic sonar system performing spectrum-based localization. IEEE Trans Robotics 23(6): 1151–1159

Skolnik M (2008) Radar handbook 3rd Ed. McGraw-Hill, New York

Smith R, Cheeseman P (1986) On the representation and estimation of spatial uncertainty. Int J Rob Res 5(4): 56–68

Suga N (1990) Cortical computational maps for auditory imaging. Neural networks 3: 3–21

Surlykke A, Ghose K, Moss C (2009) Acoustic scanning of natural scenes by echolocation in the big brown bat, *Eptesicus fuscus.* J Exp Biol 212: 1011–1020

Surlykke A, Moss CF (2000) Echolocation behavior of big brown bats, *Eptesicus fuscus*, in the field and the laboratory. J Acoust Soc Am 108: 2419–2429

Thrun S, Burgard W, Fox D (1998) A probabilistic approach to concurrent mapping and localization for mobile robots. Machine Learning and Autonomous Robots (joint issue) 31/5: 1–25

Active sensing: head and vibrissal velocity during exploratory behaviors of the rat

14

R. Blythe Towal, Brian W. Quist, Joseph H. Solomon, Mitra J. Z. Hartmann

Contents

Abstract

The vibrissal-trigeminal pathway of the rat has become an increasingly important model in neuroscience to study how sensory and motor signals are encoded, processed, and integrated in the nervous system, ultimately yielding "perception" of an object. In this chapter, we focus specifically on the role of head and vibrissa (whisker) velocity during exploratory movements. The chapter begins by describing basic vibrissal anatomy and mechanics, and shows that different studies measure "vibrissa velocity" under very different mechanical conditions, which will give rise to very different types of mechanoreceptor activation. It is thus critical to consider forces and bending moments at the whisker base in addition to vibrissa velocity when quantifying vibrissa-object contact during natural behavior. To illustrate this point, we summarize recent results demonstrating that whisking velocity at the time of collision with an object may influence the rat's ability to determine the radial distance to the object as well as the horizontal angle of contact. Further, we present evidence suggesting that the rat may actively select velocities at different points in the whisking trajectory, perhaps to aid localization behavior in these two dimensions. Finally, because the whiskers are always acting in concert with the head, we describe correlations

Mitra J. Z. Hartmann
Northwestern University
Departments of Mechanical Engineering
and Biomedical Engineering
2145 Sheridan Road, Evanston, IL 60208-3107, USA
e-mail: m-hartmann@northwestern.edu

between whisking behavior and head velocity. Preliminary data suggest that the position, orientation, and velocity of the head – which moves at a very different spatial and temporal scale than the vibrissae – will have a large effect on the tactile information acquired by the vibrissal system.

1. Introduction

Rats are nocturnal animals with low acuity vision (Dean 1981; Silveira et al. 1987; Keller et al. 2000; Prusky et al. 2000; 2002). They use rhythmic (5–25 Hz) movements of their vibrissae (whiskers) to tactually extract object features, including size (Brecht et al. 1997; Krupa et al. 2001), shape (Brecht et al. 1997; Harvey et al. 2001), orientation (Polley et al. 2005), and texture (Carvell and Simons 1995; Nicolelis et al. 1996; Arabzadeh et al. 2003; 2004; 2006; Neimark et al. 2003; Moore 2004; Hipp et al. 2006; Vaziri et al. 2007). These rhythmic vibrissal movements are known as "whisking," and provide the rat a rich tactile window to its environment.

As shown in Fig. 1A, about 30 vibrissae are arranged in a regular array on each side of the rat's face (Vincent 1913). Although there are no sensors along the length of a vibrissa, each vibrissa base is embedded within a densely innervated follicle (Fig. 1B) (Rice et al. 1986; 1993; 1997; Mosconi et al. 1993; Ebara et al. 2002). Mechanoreceptors in the follicle transduce deformations into electrical signals and provide input to primary sensory neurons in the trigeminal ganglion. (Zucker and Welker 1969; Gibson and Welker 1983a; b; Lichtenstein et al. 1990; Stüttgen et al. 2006; 2008; Szwed et al. 2006; Leiser and Moxon 2007; Khatri et al. 2009). Within nearly every brain structure of the ascending vibrissal-trigeminal system, neurons are grouped so as to directly reflect the regular peripheral topography (Woolsey et al. 1975; Van der Loos 1976; Belford and Killackey 1979; Killackey 1980; Arvidsson and Rice 1991). The presence of these neural maps allows researchers to record from neurons

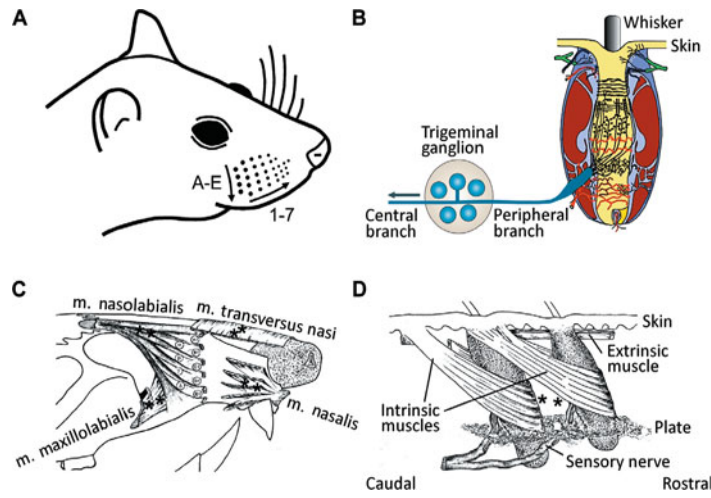


Fig. 1 Vibrissa anatomy and mechanics. **A** Vibrissae are arranged in rows and columns on the rat's face. **B** Peripheral branches of trigeminal ganglion neurons receive input from mechanoreceptors in the vibrissa follicle. Drawing adapted from Diamond et al. (2008), and Rice et al. (1997). **C** Four extrinsic muscles attach the mystacial pad to the skull. **D** The intrinsic muscles form "slings" around the follicles, and join adjacent follicles of a single row. Drawings in C and D were adapted from Hill et al. (2008), which were in turn adapted from Fig. 1 and 3 of Dörfl (1982; 1985)

responsive to particular groups of vibrissae, and to study how tactile information is integrated with motor signals (Ahissar and Arieli 2001; Kleinfeld et al. 2006; Diamond et al. 2008). This makes the rat vibrissal system an excellent model to explore active sensing behaviors and the structure of movements that subserve sensing.

Given the importance of this model system to neuroscience, it is somewhat remarkable that we have not yet fully identified the mechanical parameters at the vibrissa base relevant to behavior, and how these parameters are encoded in the earliest stages of the vibrissal-trigeminal pathway. Until the mechanical input to the vibrissal pathway is quantified, our ability to interpret the functional and computational characteristics of higher-level neural processing stages will remain severely limited.

One reason that it has historically been difficult to quantify the mechanical variables important to whisking behavior is that whisking is fundamentally a process of "active sensing," in other words, the mechanical signals obtained by the vibrissae depend on how they are moved. In principle, the rat could – and probably does – use any number of control strategies for exploration with vibrissae. For example, it could control the angle of its vibrissae, the velocity of protractions and retractions, or the force required to move its vibrissae through a particular angle. In this chapter, we focus specifically on the role of head and whisking velocity during exploratory movements.

This chapter consists of three main parts. First, we describe basic vibrissa anatomy and mechanics. An important point in this section is that different studies measure "vibrissa velocity" under very different mechanical conditions. Second, we show that whisking velocity at the time of collision with an object will influence the rat's ability to determine both radial contact distance as well horizontal contact angle. We present evidence suggesting that the rat may tune or select velocities at different points in the whisking trajectory, perhaps to aid in localization behavior. Finally, we describe correlations between whisking behavior and head velocity, and suggest some ways in which coordinated head and vibrissa movements may enhance sensing.

2. Vibrissa mechanics: active movements vs. passive stimulation

2.1 Vibrissa and follicle anatomy and muscle mechanics

Rat vibrissae have an intrinsic curvature (Knutsen et al. 2008; Towal et al. 2011) and taper approximately linearly to a diameter of a few micrometers at the tip (Williams and Kramer 2010). The proximal portion of the vibrissa (~60–70%) typically lies in a plane (Knutsen et al. 2008; Towal et al. 2011). That is, if one were to pluck the vibrissa out and place it on a table, the proximal portion would lie flat, in the plane of the table. The remaining fraction of the vibrissa generally curves out of the plane.

The base of each vibrissa inserts into a follicle replete with mechanoreceptors. The responses of these mechanoreceptors to a particular mechanical input will be determined in part by the stiffness of the surrounding tissue. In addition, a large blood sinus occupies much of the follicle, and its degree of engorgement with blood may modulate the stiffness with which the vibrissa is held. Finally, the musculature surrounding the follicle is also likely to alter the stiffness with which the vibrissa is held at its base and with which it resists deflection.

The vibrissae are actuated by two dis-

tinct groups of muscles: extrinsic muscles connect the mystacial pad directly to the skull, while intrinsic muscles form a "sling" around each of the follicles (Dörfl 1982; 1985, see Fig. 1C–D). The intrinsic and extrinsic muscles act in a three-step sequence during the whisk cycle (Dörfl 1982; 1985; Wineski et al. 1988; Berg and Kleinfeld 2003; Hill et al. 2008). First, a rostral extrinsic muscle (*m. nasalis*) initiates protraction by pulling the mystacial pad forward. Second, the intrinsic sling muscles contract around each follicle to rotate the vibrissa farther forward. Finally, two large caudal muscles (*m. nasolabialis* and *m. maxiolabialis*) contract to pull the mystacial pad and vibrissae back to their initial state, aided by the elasticity of the skin.

In the head-restrained or over-trained animal, whisking trajectories can appear quite stereotyped (Bermejo et al. 2002; Mehta et al. 2007). Under more natural conditions, however, involving head movements or contact with objects, the vibrissae can exhibit considerable variability in their kinematic profiles (Wineski 1983; Sachdev et al. 2003; Towal and Hartmann 2006; 2008; Mitchinson et al. 2007; Grant et al. 2009). The trajectories are even more complex if the very tip of the vibrissa is considered. Because the vibrissae are flexible and actuated from the base, tip velocities involve substantial dynamics. Accordingly, vibrissa "velocity" is typically used to refer only to movements of the proximal portion of the vibrissa as it emerges from the mystacial pad. To first approximation, this portion of the vibrissa can be taken to move as a rigid body during free-air whisking (Knutsen et al. 2008), but not during collisions or periods of contact with an object.

A recent study showed that as the vibrissa protracts forward, it exhibits substantial "roll" about its long axis (Knutsen et al. 2008). This same study also showed that over the course of the whisking cycle, the roll angle, ζ, correlates strongly with horizontal angle, θ, (Knutsen et al. 2008). When the head is static, the orientation of the vibrissa's intrinsic curvature relative to an object is determined by the roll angle. Because the roll angle varies with the horizontal angle, the vibrissa will collide with an object at different orientations, depending on the horizontal angle at which the collision occurs. When the head is in motion, both the roll angle and the orientation of the head relative to the object will determine the orientation at which the vibrissa collides with the object. Thus, at the time of a collision, the vibrissa may have its concave side facing the object, its convex side facing the object, or anything in between.

The intrinsic curvature of the vibrissa as it collides with an object will in turn affect the forces generated during collision. For example, collisions with the concave side of the whisker will generate a larger net force vector and longer duration contacts than collisions with the convex side of the whisker (Quist and Hartmann, submitted). This in turn leads to the behavioral prediction that the rat might sometimes alter its exploratory strategies to ensure collisions with the vibrissa's concave side.

The forces generated by a collision are mechanically transmitted by the vibrissa to its base, where, as described above, mechanoreceptors in the follicle transduce mechanical deformations to electrical signals. Deformation of a mechanoreceptor could result if the vibrissa moves relative to the follicle; it could result from muscles (either intrinsic or extrinsic) squeezing on the outside of the follicle; or it could result from the blood sinus distending or relaxing. These complex features of mechanical transduction in the follicle mean that boundary conditions and the shape of the vibrissa at contact will clearly have a large effect on the incoming sensory information.

2.2 Different studies measure "vibrissa velocity" under very different mechanical conditions

All studies to date have measured "vibrissa velocity" in essentially the same way, but, as will be shown, under very different mechanical conditions. Vibrissa angle is measured near the base, along the initial linear portion of the vibrissa (θ in Fig. 2A). The first time derivative of the angle is then taken as the velocity. The apparent uniformity across studies in the method for measuring "vibrissa velocity," however, belies the very different mechanical conditions under which this quantity is measured.

Many studies of whisking behavior in the awake animal measure kinematics during

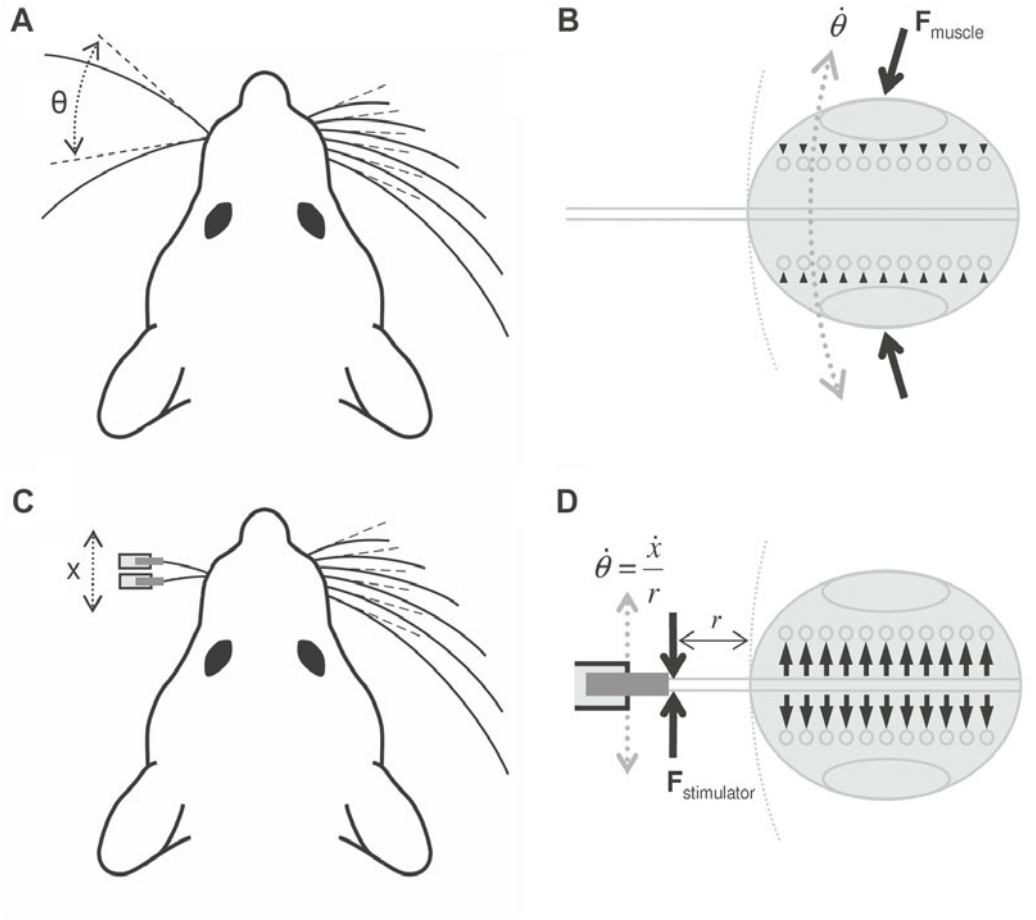


Fig. 2 Vibrissa velocity measured during free-air whisking and during passive vibrissa displacements. **A** Overhead view of rat whisking in free air. A single vibrissa on the left mystacial pad is shown to sweep through the angle θ. **B** Overhead view of passive deflection of a vibrissa by a stimulator. Because displacements are typically small, they are measured linearly (x). **C** During free-air whisking behavior, forces (large solid black arrows) from the surrounding musculature act on the follicle-sinus-complex as a unit. Forces acting on the mechanoreceptors (stylized as gray circles) are likely to be small. **D** During passive stimulation, an external force is imposed directly on the vibrissa shaft and transmitted into the follicle. The vibrissa will exert forces (solid black arrows) directly on the mechanoreceptors (stylized as gray circles) and a large response will be generated

free-air whisking (Sachdev et al. 2002; Berg and Kleinfeld 2003; Jin et al. 2004; Towal and Hartmann 2006; 2008; Khatri et al., 2009). During free-air whisking, the intrinsic and extrinsic muscles move the follicle and mystacial pad as described earlier, resulting in the type of motion schematized in Fig. 2A. Protraction velocities in these studies typically range from 250°/sec to 1000°/sec; retraction velocities typically have a higher upper bound, near 1500°/sec. Slightly lower values (~700°/sec and ~1100°/sec for retraction) were found for protractions and retractions over a textured surface (Carvell and Simons 1990).

In contrast, studies describing neural recordings from the anesthetized animal generally measure velocity of the vibrissa when it is rigidly attached to a stimulator and passively displaced (Pinto et al. 2000; Shoykhet et al. 2000; Jones et al. 2004). The stimulator is often a piezoelectric crystal, a stepper motor, or a solenoid, and is moved either in a "ramp-and-hold" paradigm or oscillated at known frequencies. Because these passive displacements are typically small, and the stimulator is often translated instead of rotated, velocities are often expressed as linear measures (mm/sec). This results in the type of motion shown in Fig. 2B. When converted to angular velocities, the linear velocities in these studies typically range from 700°/sec to 2000°/sec.

These two types of experimental conditions could produce *identical* "velocities," but they are mechanically very different, and will generate very different deformations of mechanoreceptors in the follicle. The differences become apparent when considering the anatomy of the vibrissa/follicle complex together with the locations of the receptors, as shown in Fig. 2C–D.

In the case of free air whisking, the vibrissa and the follicle are actuated together, as a unit, by the muscles surrounding the follicle. This is the situation depicted in Fig. 2C, in which two black solid arrows illustrate the force of the surrounding muscles and skin tissue on the follicle. In this case, there is likely to be minimal relative motion between the vibrissa and follicle. If there is little or no relative motion, mechanoreceptors in the follicle will deform only slightly, and only a very weak signal will be transmitted to primary sensory neurons. This is consistent with recent findings that kinematic variables (deflection amplitude, velocity, position) are poorly coded by primary sensory neurons during free-air whisking behavior (Khatri et al. 2009). The small correlation that does exist may be due to the muscles squeezing down on the follicle or inertial effects deforming the vibrissa within the follicle (small black arrowheads in figure).

If instead a stimulator is used to passively displace the vibrissa, as during experiments with anesthetized animals, the applied force will generate forces and moments via the vibrissa shaft *in* the follicle (Pinto et al. 2000; Shoykhet et al. 2000; Jones et al. 2004). This is illustrated in Fig. 2D by the solid black arrows, now on the *inside* of the follicle, depicting the vibrissa pressing against a set of mechanoreceptors. Direct activation of mechanoreceptors during passive stimulation is consistent with the strong responses of primary ganglion neurons in these experiments (Gibson and Welker 1983a; b; Lichtenstein et al. 1990; Leiser and Moxon 2006; Szwed et al. 2006).

It is important to note that we cannot determine the precise distribution of forces on mechanoreceptors in the follicle, because we do not yet know exactly how stiffly (or loosely) the vibrissa is held at every point within the follicle. Regardless of the precise distribution of forces, however, it is clear that during free-air whisking the driving force originates in the muscles outside of the follicle, while during passive stimulation, the driving force originates from the vibrissa shaft inside of the follicle.

2.3 Interpretation of studies using ramp-and-hold stimuli and oscillatory stimuli

To summarize so far, the velocity of a vibrissa does not uniquely determine the forces and moments generated at its base. Muscles may pull on the vibrissa-follicle complex to generate a velocity of 1000°/second, but if there is no relative motion between vibrissa and follicle, mechanoreceptors will not deform. A situation similar to this may arise as the rat whisks freely in air. Conversely, if an object near the vibrissa base obstructs the vibrissa's movement, its velocity might approach zero, but generate forces and moments that cause substantial mechanoreceptor deformation.

Comparing "vibrissa velocity" during passive stimulation to "vibrissa velocity" during free air whisking is like comparing apples and oranges. Studies that employ passive ramp-and-hold stimuli in anesthetized animals are most likely examining how neurons respond to rates of change of forces and moments, as might occur during collision with an object.

For example, previous studies using ramp-and-hold stimuli have shown that rat barrel cortical neurons respond preferentially to "high-velocity" ramp stimuli. These stimuli were found to produce highly synchronized responses in the ventral posterior medial thalamus (VPm), which in turn produced strong responses in layer 4 barrel cortex (Pinto et al. 1996; 2000; Temereanca and Simons 2003). In contrast, large amplitude movements at low velocities did not evoke strong responses in the cortex despite causing high thalamic discharge rates (Pinto et al. 2000). These studies concluded that the strength of cortical responses (in number of action potentials and duration of response) was directly proportional to vibrissa velocity. But from a behavioral point of view, these neurons would respond most strongly to rapid force and moment changes associated with collisions with objects. The neurons would be unlikely to respond to high velocities during free-air whisking behavior.

Similarly, studies that rapidly oscillate the vibrissa back and forth about its rest point are inducing rapid changes in forces and moment that do not resemble those that would be generated either by free-air whisking, or by collision with an object. It is difficult to say what behavioral condition these experiments replicate. The oscillations generated may be similar to the rapid vibrations that could be induced as a rat swept its vibrissae over a texture, but they are not quite the same, because the vibrissa is being driven quite differently than would occur during natural texture exploration (Hartmann et al. 2003; Neimark et al. 2003).

In the next section, we discuss how vibrissa velocity influences the forces and moments that will be generated at the vibrissa base, and the roles that velocity may play in the rat's exploratory strategies.

3. The importance of velocity in object localization

The natural exploratory behavior of the rat is characterized by "bouts" of whisking that last between a few hundred milliseconds up to a few seconds (Welker 1964; Berg and Kleinfeld 2003). During tactual exploration of the environment, the rat moves its body and head so that its vibrissae are sometimes moving in free air, but at other times colliding with various objects in the environment.

When the vibrissa collides with an object, any one of several behavioral events can occur. One possibility is that the rat may bring

its vibrissa to a very quick halt and then reverse its direction of motion (Grant et al. 2009). On the opposite extreme, the rat may continue to brush the vibrissa along the object, so that at some point the vibrissa may ultimately slip past it and lose contact during protraction (Hartmann et al. 2003). Regardless of which behavior occurs, there are three important points to be made about the collision process in the context of object localization.

First, contact with an object can occur at any location along the length of the vibrissa. Contact does not need to occur at the vibrissa tip. This leaves the rat with the problem of determining both the radial distance to the object and the angle of contact. Somehow, the forces and moments induced by the collision must be interpreted by the nervous system to yield an impression of a spatial coordinate in (*r*, *θ*, and *z*). This problem is illustrated in Fig. 3A. Note that the height (*z*) of a given vibrissa is thought to be coded by a labeled line strategy (Diamond et al. 2008)

Second, mechanoreceptor deformation is the only event that can let the animal know that contact with an object has occurred. Receptor deformation will be induced by the forces and moments (including vibrations) generated by collisions. This is because the collision generates reaction-forces and reaction-moments at the vibrissa base, as shown in the free-body diagram of Fig. 3B. Ultimately, forces and moments are the physical variables corresponding to "touch." Work from multiple laboratories has convincingly demonstrated that position or phase information must be combined with "touch" information in order for the rat to determine the horizontal angle of contact (Szwed et al. 2006; Mehta et al. 2007; Diamond et al. 2008).

Third, the particular forces and moments generated by the collision will depend on the vibrissa velocity at the time of collision. In the next section we describe how the velocity at the time of collision may affect how the rat localizes the horizontal angle of the object. We then present evidence to suggest that the rat may control or "tune" the instantaneous whisking velocity during the whisk cycle, perhaps to aid in this process of localization.

3.1 Role of velocity in determining object coordinates: radial distance and horizontal angle

Our laboratory recently described a mechanism through which the rat could determine the radial distance from the vibrissa base to an object. (Solomon and Hartmann 2006;

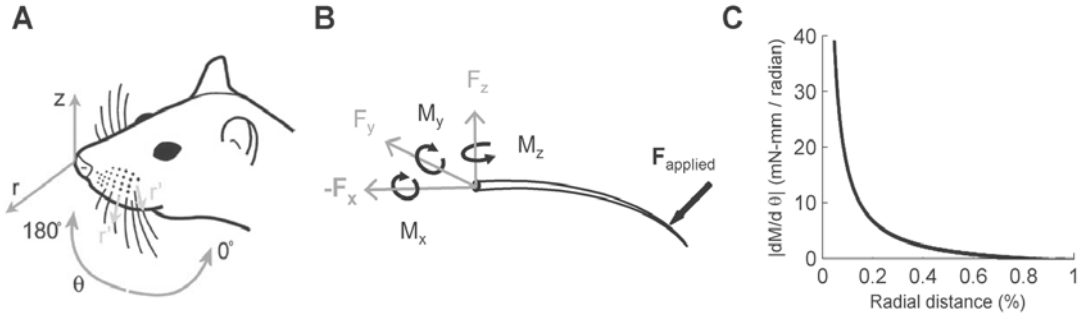


Fig. 3 **A** The rat must localize an object in 3-dimensions relative to the snout: *r, θ*, and *z*. Fig. adapted from Gopal and Hartmann (2007). **B** Free-body diagram of vibrissa. **C** *dM*/*dθ* as a function of normalized contact distance. The plot shows simulation results for a straight, linearly tapered vibrissa deflecting against a frictionless peg

Birdwell et al. 2007). Our method was based on an approach first described by Kaneko in 1998 (Kaneko et al. 1998). When the vibrissa collides with an object, the resulting force causes the vibrissa to bend. The amount of bending is proportional to the moment M generated at the base, and the quantity $dM/d\theta$ can be shown to be monotonically related to the radial object distance, r, as shown in Fig. 3C (Kaneko et al. 1998; Birdwell et al. 2007). This means that if the rat can keep track of the rate of change of moment (dM/dt) and the velocity ($d\theta/dt$) with which it is "pushing" its vibrissa against the object, enough information will be present to determine $dM/d\theta$ and thus the radial object distance *r.* In the same set of papers, we also noted that the rat could choose to "tune" the particular value of dM/dt associated with a particular value of *r,* simply by changing the velocity of its whisk. If the rat whisks at a higher velocity, the rate of moment change dM/dt will also be higher; the rat could potentially tune its whisking velocity to bring dM/dt into the most sensitive (highest resolution) region.

Velocity may be similarly important in determining the horizontal angular location of an object, the variable θ in Fig. 3A. As discussed above, in order to determine the angular location of an object in the horizontal plane, the rat must combine kinematic information with information about the time or location of object collision. There are currently two hypotheses for the computation of horizontal location.

One hypothesis predicts that the nervous system performs coincidence detection between the responses of *position*-sensitive neurons and contact-sensitive neurons (Szwed et al. 2003; Mehta et al. 2007) to determine the horizontal angle of an object. A second hypothesis, however, predicts that the nervous system performs coincidence detection between the responses of *phase*-sensitive neurons and contact-sensitive neurons to determine the horizontal angle of an object (Ahissar 1998; Ahissar and Arieli 2001; Szwed et al. 2003; Curtis and Kleinfeld 2009). Notice that the definition of phase used in these studies is a *temporal* phase, found by dividing each whisk cycle evenly between 0 and 2π radians. This hypothesis requires that the velocity of the protraction always have the same sign (i. e., that the vibrissa does not reverse direction during a protraction). Otherwise, the relationship between temporal phase and position is not unique, and horizontal location cannot be determined. Velocity is thus likely to play a critical role in spatial localization in both the radial and horizontal dimensions. In the next section, we present evidence that the rat may tune instantaneous whisking velocity during the whisk cycle, perhaps to aid in this process of localization.

3.2 Variations in whisking velocity during the whisk cycle

We recently quantified variability in the velocity profiles of whisking during natural exploratory behavior (Towal and Hartmann 2008). The results showed that although most retractions consisted of smooth, monotonic velocity profiles, the majority of protractions were not smooth or monotonic. Most notably, a significant number of the protractions showed a slowing mid-trajectory, and many actually reversed direction during the course of the protraction. Examples of these three types of whisking profiles are illustrated for protractions in Fig. 4. It is clear from the figure that different whisking velocity profiles achieve maximum velocity at different temporal phases of the whisk.

Importantly, however, the results also indicated that despite the high degree of variability in the instantaneous whisking

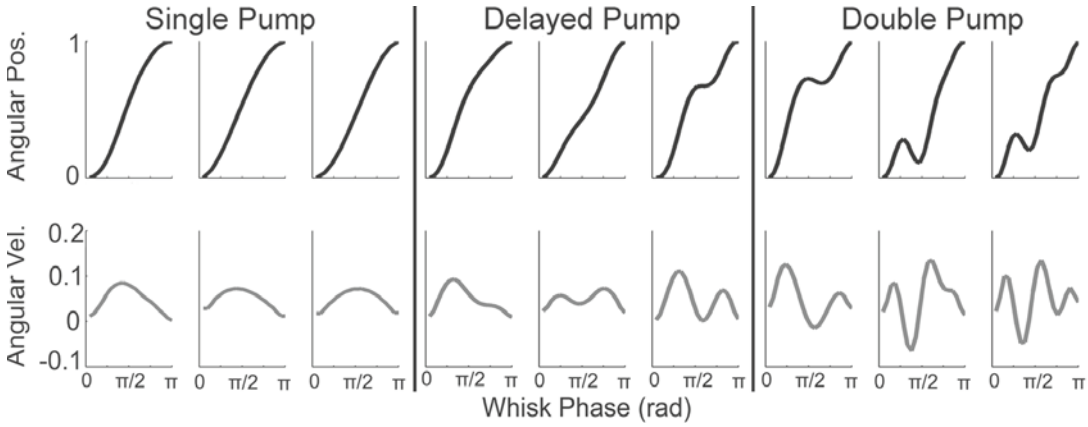


Fig. 4 Nine exemplary protraction segments illustrate the continuum of position and velocity profiles. Angular position has been normalized between zero and one. Angular velocity was calculated as the derivative of the normalized angular position. Both velocity and position are plotted as a function of temporal phase, with π indicating full protraction. Thick black vertical lines indicate the separation of the continuum into the three profile types: single, delayed, and double

velocity, the average whisking velocity remained remarkably constant from whisk to whisk. This constancy was possible because different whisking profiles appeared to compensate for any initial "error" in the velocity of the whisk segment. If a whisk segment started faster than average, then a delayed or double pump whisk profile subsequently slowed it down. If a whisk started more slowly than average, then the rat tended not to execute a double or delayed pump, but rather to execute a single pump that increased in speed.

Thus, the primary result of the study was that each protraction and retraction may exhibit a large variability in instantaneous velocity but – across whisks – demonstrate a low variability in average whisking velocity. This suggests that the rat may be exploiting the tri-phasic muscle activation pattern to actively adjust its whisking velocity in real time. In principle, these velocity adjustments may alter the rate of moment change so as to improve the extraction of radial distance. For example, the rat may want to increase the rate of moment change for contact with objects near the tip to increase signal size.

In addition, these results argue against the temporal-phase hypothesis for horizontal object localization, because the relationship between whisk phase and external space is unique only when the vibrissa position monotonically increases through the whisk. If a double pump is generated by a single triphasic pattern of muscle activity, then a reversal in direction occurs during the whisk. This means that the relationship between temporal phase and external space is non-unique, as the same spatial location is achieved at three distinct times (temporal phases) during the whisk. Instead, our data favor the position-hypothesis for determination of the horizontal contact angle, consistent with recent findings from the Kleinfeld laboratory (Mehta et al. 2007). By directly measuring vibrissa position, the delayed and double pumps no longer pose a problem for calculating the spatial position of contact.

4. The relation between head velocity and whisking movements

As described in section 2.1, the velocity of the vibrissa as it makes contact with an object will determine the changes in the forces and moments at the vibrissa base. When the head is static, the rat can control vibrissa velocity via its intrinsic and extrinsic muscles (Fig. 1). If the head is free to move, however, then the velocity of the vibrissa relative to the object also depends on the head velocity, because the vibrissa is attached to the head. In this section we describe the coupling that exists between head and vibrissa movements.

4.1 Right-left positional asymmetries are correlated with rotational head velocity

We recently performed a study to examine the extent to which bilateral free-air whisking movements were influenced by head rotations in the horizontal plane. We initially hypothesized that the rat might want to maintain spatial and temporal symmetry in world coordinates. That is to say, the vibrissae should maintain their velocity in the world reference frame *regardless* of the head velocity. For example, suppose that the rat is turning its head to the right with a velocity ω_{head}. In this case, the right vibrissae would need to speed up by a velocity exactly equal to ω_{head} and the left vibrissae would need to slow down by a velocity ω_{head}. This would mean, in turn, that the difference of left and right vibrissa velocities should be proportional to twice the head velocity. (In practice, this difference becomes a sum because left vibrissa velocities are neg-

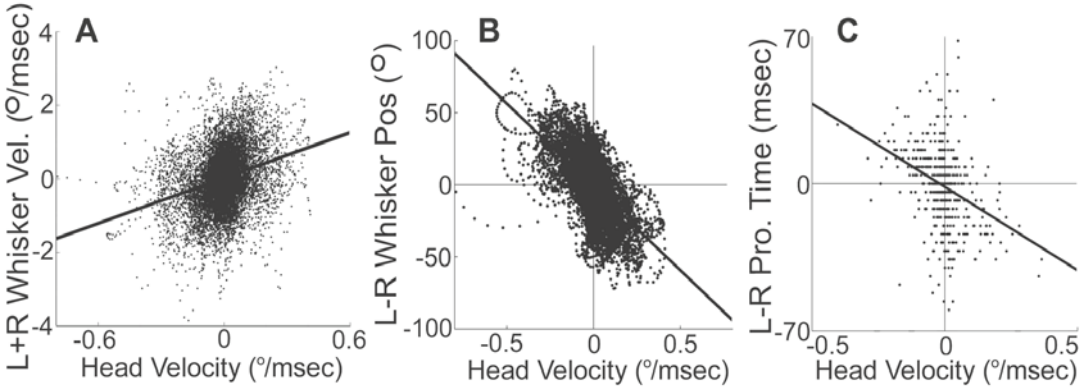


Fig. 5 Relationships between whisking parameters and head velocity. In all graphs, head velocities are represented in world coordinates, with negative velocity indicating that the head is turning to the right. **A** Velocity asymmetry does not strongly depend on head velocity. In world coordinates, the sum of whisking velocities correlates weakly with head velocity with a slope of 2.06. This slope is statistically indistinguishable from the slope of 2.00 that would occur if one side sped up by the head velocity, while the other side slowed down by the head velocity ($p = 0.30$). **B** Spatial asymmetry does depend on head velocity. Instantaneous left–right vibrissa position difference versus instantaneous head velocity plotted continuously over 373 whisks. The slope of the best linear fit is 115 msec, which approximately equals the average duration of a single whisk (121 msec). The r-value is 0.58. **C** Temporal asynchrony is not well correlated with head velocity The largest correlation ($r = 0.34$) between left–right vibrissa time differences and head velocity was found for protractions, with the head velocity averaged over 52 msec following the protraction

ative when measured in world coordinates.) The results of this analysis are shown in Fig. 5A. Although the slope is very close to 2, as predicted, the correlation is weak, with an r-value of 0.28. This led us to reject our initial hypothesis.

Instead, as shown Fig. 5B, left-right whisking asymmetry was found to be strongly correlated with rotational head velocity. Specifically, the positional asymmetry of the left and right vibrissae arrays was equal to the distance that the head would rotate during a whisk. In other words, the vibrissae "look ahead" of the current position of the head to anticipate head movement that would occur, were the rat to continue to rotate its head at that velocity. The rat thus appears to use its vibrissae to search in advance of the head, presumably to avoid running into unexpected obstacles.

Surprisingly, however, there was only a weak relationship between head velocity and right–left time differences of peak protraction and retraction. Fig. 5C shows that although temporal asynchrony is large (up to 70 msec), it is not nearly as well correlated with head velocity as is spatial asymmetry. Temporal asynchronies could, theoretically, be used to differentially search particular regions of space during head rotations. For example, the rat could choose to stop protraction on right and left sides at different points in time, thus selecting whisk amplitudes and the particular (spatial) regions swept out by each side.

In summary, during free air whisking behavior that includes head rotations, the rat is not trying to ensure that right and left vibrissa arrays have the same velocity, or that right and left vibrissae arrays are in the same phase of the whisk. Neither are the protraction and retraction times strongly related to the rotational head velocity. Instead, the angular positions of left and right arrays are strongly correlated with the velocity of the head.

4.2 Future directions: Simultaneous measurement of head and vibrissa velocities during contact with an object

Until recently, the small size and rapid speed of vibrissa movements precluded quanti-

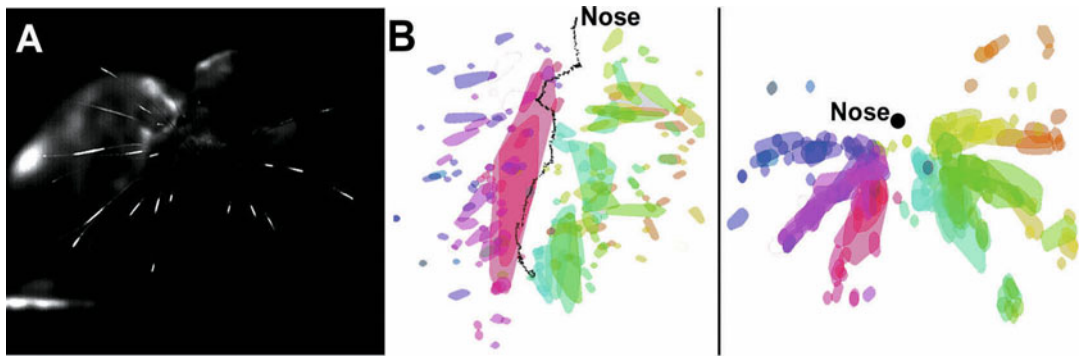


Fig. 6 Vibrissa-object contact patterns measured with a laser light sheet. **A** Image from a high speed camera obtained as the rat explored a flat glass plate with the light sheet in front. The points of contact that the vibrissa makes with the sheet are seen as white dots. **B** Contact patterns of vibrissae on the glass plate over a period of three whisks (~360 msec) *Left:* Spatial distribution of vibrissa contacts as the rat's head moved in an approximately-vertical trajectory (dark line labeled "nose"). The contacts of each vibrissa are represented in a different color. *Right:* The same data as in the left plot, after subtracting out head movement. A high degree of spatial overlap across the three whisks is revealed, demonstrating that head movement accounts for the large spatial coverage in the left plot

fication of the patterns of contact that the vibrissae make with an object, and there was no method to examine how head and vibrissa velocities might change during the exploration process. Our laboratory has recently developed a technique to permit visualization of vibrissa-object contact patterns during the rat's natural exploratory behavior (Towal and Hartmann 2010). Light from a laser is passed through a set of lenses to create a thin, planar light sheet. The light sheet cascades vertically down immediately in front of a flat glass wall. Fig. 6A shows one frame from a high speed video camera (1000 fps) obtained as a rat explored the flat glass surface. Vibrissae in contact with the glass are clearly seen as white dashes. As shown in Fig. 6B, we have used this technology to quantify the external space contacted by each vibrissa as the rat explores the glass sheet. It is clear from the figure that head movements account for much of the variability in the spatial locations of vibrissa-object contact. When head movements are removed, the vibrissae remain primarily in narrow "slots," associated with their position in the array. This was confirmed using step-wise regression techniques which indicate that both whisker and head movement variables are required to explain a significant amount of the variability in the whisker-object contact patterns.

Conclusions

This chapter has three main conclusions.

First, and most importantly: different studies measure "vibrissa velocity" under extremely different mechanical conditions. Passively shaking a vibrissa at 8 Hz, for example, does not replicate the mechanical conditions associated with 8 Hz free-air whisking.

Second, the velocity of whisking is likely to be an important control parameter for the rat. The rat could vary whisking velocity to improve the accuracy of radial distance determination.

Third, the position, orientation, and velocity of the head will have a large effect on the tactile information acquired by the vibrissae. Like the eyes and the fingers, the vibrissae are relatively low mass, high resolution sensing structures, whose movements may be controlled nearly kinematically. Eyes, vibrissae and fingers are all located on far more massive structures (the head and the forelimb) whose mass cannot be neglected. Movements of the head and forelimb are generally much slower than those of the vibrissae, eyes, and fingers, and they serve to place these fine-resolution sensory structures on the portion of the surface that the animal will explore next. Future work will aim to quantify the relative contributions of head and vibrissa movements – operating at different spatial and temporal scales – to the acquisition of sensory information as the rat explores an unknown environment.

References

Ahissar E (1998) Temporal-code to rate-code conversion by neuronal phase-locked loops. Neural Comput 10: 597–650

Ahissar E, Arieli A (2001) Figuring space by time. Neuron 32: 185–201

Arabzadeh E, Panzeri S, Diamond ME (2004) Vibrissa vibration information carried by rat barrel cortex neurons. J Neurosci 24: 6011–20

Arabzadeh E, Panzeri S, Diamond ME (2006) Deciphering the spike train of a sensory neuron: Counts and temporal patterns in the rat vibrissa pathway. J Neurosci 26: 9216–26

Arabzadeh E, Petersen RS, Diamond ME (2003) Encoding of vibrissa vibration by rat barrel cortex neurons: Implications for texture discrimination. J Neurosci 23: 9146–54

Arvidsson J, Rice FL (1991) Central projections of primary sensory neurons innervating different parts of the vibrissae follicles and intervibrissal skin on the mystacial pad of the rat. J Comp Neurol 309: 1–16

Belford GR, Killackey HP (1979) Development of vibrissae representation in sub-cortical trigeminal centers of the neonatal rat. J Comp Neurol 188: 63–74

Berg RW, Kleinfeld D (2003) Rhythmic whisking by the rat: Retraction as well as protraction of the vibrissae is under active muscular control. J Neurophysiol 89: 104–17

Bermejo R, Vyas A, Zeigler H (2002) Topography of rodent whisking. I. Two dimensional monitoring of vibrissa movements. Somatosens Mot Res 19: 341–6

Birdwell JA, Solomon JH, Thajchayapong M, Taylor M, Cheely M, Towal RB, Conradt J, Hartmann MJZ (2007) Biomechanical models for radial distance determination by rat vibrissae. J Neurophysiol 98: 2439–55

Brecht M, Preilowski B, Merzenich MM (1997) Functional architecture of the mystacial vibrissae. Behav Brain Res 84: 81–97

Carvell GE, Simons DJ (1990) Biometric analyses of vibrissal tactile discrimination in the rat. J Neurosci 10: 2638–48

Carvell GE, Simons DJ (1995) Task- and subject-related differences in sensorimotor behavior during active touch. Somatosens Mot Res 12: 1–9

Dean P (1981) Visual pathways and acuity in hooded rats. Behav Brain Res 3: 239–71

Diamond ME, von Heimendahl M, Knutsen PM, Kleinfeld D, Ahissar E (2008) 'Where' and 'what' in the vibrissa sensorimotor system. Nature Rev Neurosci 9: 601–12

Dörfl J (1982) The musculature of the mystacial vibrissae of the white mouse. J Anat 135: 147–54

Dörfl J (1985) The innervation of the mystacial region of the white mouse. A topographical study. J Anat 142: 173–84

Ebara S, Kumamoto K, Matsuura T, Mazurkiewicz JE, Rice FL (2002) Similarities and differences in the innervation of mystacial vibrissal follicle-sinus complexes in the rat and cat: A confocal microscopic study. J Comp Neurol 449: 103–19

Gibson JM, Welker WI (1983a) Quantitative studies of stimulus coding in first-order vibrissa afferents of rats. 1. Receptive field properties and threshold distributions. Somatosens Res 1: 51–67

Gibson JM, Welker WI (1983b) Quantitative studies of stimulus coding in first -order vibrissa afferents of rats. 2. Adaptation and coding of stimulus parameters. Somatosens Res 1: 95–117

Gopal V, Hartmann MJZ (2007) Using hardware models to quantify sensory data acquisition across the rat vibrissal array. J Bioinsp Biomim 2: 135–145

Grant RA, Mitchinson B, Fox CW, Prescott TJ (2009) Active touch sensing in the rat: anticipatory and regulatory control of vibrissa movements during surface exploration. J Neurophysiol 101: 862–74

Hartmann MJ, Johnson NJ, Towal RB, Assad C (2003) Natural resonance frequencies and damping characteristics of rat vibrissae. J Neurosci 23: 6510–6519

Harvey MA, Bermejo R, Zeigler HP (2001) Discriminative whisking in the head-fixed rat: optoelectronic monitoring during tactile detection and discrimination tasks. Somatosens Mot Res 18: 211–22

Hill, DN Bermejo R, Zeigler HP, Kleinfeld D (2008) Biomechanics of the vibrissa motor plant in rat: Rhythmic whisking consists of triphasic neuromuscular activity. J Neurosci 28: 3438–55

Hipp J, Arabzadeh E, Zorzin, E, Conradt J, Kayser C, Diamond ME, Konig P (2006) Texture signals in vibrissa vibrations. J Neurophysiol 95: 1792–9

Jin TE, Witzemann V, Brecht M (2004) Fiber types of the intrinsic vibrissa muscle and whisking behavior. J Neurosci 24: 3386–3393

Jones LM, Depireux DA, Simons DJ, Keller A (2004) Robust temporal coding in the trigeminal system. Science 304: 1986–9

Kaneko M, Kanayama N, Tsuji T (1998) Active antenna for contact sensing. IEEE T Robotic Autom 14: 278–91

Keller J, Strasburger H, Cerutti DT, Sabel BA (2000) Assessing spatial vision – automated measurement of the contrast-sensitivity function in the hooded rat. J Neurosci Meth. 97: 103–10

Khatri V, Bermejo R, Brumberg JC, Keller A, Zeigler HP (2009) Whisking in air: encoding of kinematics by trigeminal ganglion neurons in awake rats. J Neurophysiol 101: 1836–46

Killackey HP (1980) Pattern-formation in the trigeminal system of the rat. Trends Neurosci 3: 303–6

Kleinfeld D, Ahissar E, Diamond ME (2006) Active sensation: insights from the rodent vibrissa sensorimotor system. Curr Opin Neurobiol 16: 435–44

Knutsen P, Biess A, Ahissar E (2008) Vibrissal kinematics in 3D: Tight coupling of azimuth, elevation, and torsion across different whisking modes. Neuron 59: 35–42.

Krupa D, Matell M, Brisben A, Oliveira L, Nicolelis MAL (2001) Behavioral properties of the trigeminal somatosensory system in rats performing vibrissa-dependent tactile discriminations. J Neurosci 21: 5752–5763

Leiser SC, Moxon KA (2006) Relationship between physiological response type (RA and SA) and vibrissal receptive field of neurons within the rat trigeminal ganglion. J Neurophysiol 95: 3129–45

Leiser SC, Moxon KA (2007) Responses of trigeminal ganglion neurons during natural whisking behaviors in the awake rat. Neuron 53: 117–33

Lichtenstein SH, Carvell GE, Simons DJ (1990) Responses of rat trigeminal ganglion neurons to movements of vibrissae in different directions. Somatosens Mot Res 7: 47–65

Mehta S, Whitmer D, Figueroa R, Williams B, Kleinfeld D (2007) Active spatial perception in the vibrissa scanning sensorimotor system. PLoS Biol 5: e15

Mitchinson B, Martin CJ, Grant RA, Prescott TJ (2007) Feedback control in active sensing: rat exploratory whisking is modulated by environmental contact. Proc R Soc Lond B 274: 1035–41

Moore CI (2004) Frequency-dependent processing in the vibrissa sensory system. J Neurophysiol 91: 2390–9

Mosconi TM, Rice FL, Song MJ (1993) Sensory innervation in the inner conical body of the vibrissal follicle-sinus complex of the rat. J Comp Neurol 328: 232–51

Neimark MA, Andermann ML, Hopfield JJ, Moore CI (2003) Vibrissa resonance as a transduction mechanism for tactile encoding. J Neurosci 23: 6499–509

Nicolelis MAL, DeOliveira LMO, Lin RCS, Chapin JK (1996) Active tactile exploration influences the functional maturation of the somatosensory system. J Neurophysiol 75: 2192–6

Pinto DJ, Brumberg JC, Simons DJ (2000) Circuit dynamics and coding strategies in rodent somatosensory cortex. J Neurophysiol 83: 1158–66

Pinto DJ, Brumberg JC, Simons DJ, Ermentrout GB (1996) A quantitative population model of vibrissa barrels: Re-examining the Wilson-Cowan equations. J Comput Neurosci 3: 247–64

Polley DB, Rickert JL, Frostig RD (2005) Vibrissa-based discrimination of object orientation determined with a rapid training paradigm. Neurobiol Learn Mem 83: 134–42

Prusky GT, Harker KT, Douglas RM, Whishaw IQ (2002) Variation in visual acuity within pigmented, and between pigmented and albino rat strains. Behav Brain Res 136: 339–48

Prusky GT, West PWR, Douglas RM (2000) Behavioral assessment of visual acuity in mice and rats. Vision Res 40: 2201–9

Quist BW, Hartmann MJZ. (2009) Experimental validation of a quasi-static model of vibrissa bending: Effects of taper and curvature. Soc Neurosci Ann Mtg, Program Number 174.6. Chicago, IL

Rice FL (1993) Structure, vascularization, and innervation of the mystacial pad of the rat as revealed by the lectin *griffonia simplicifolia*. J Comp Neurol 337: 386–99

Rice FL, Fundin BT, Arvidsson J, Aldskogius H, Johansson O (1997) Comprehensive immunofluorescence and lectin binding analysis of vibrissal follicle sinus complex innervation in the mystacial pad of the rat. J Comp Neurol 385: 149–84

Rice FL, Mance A, Munger BL (1986) A comparative light microscopic analysis of the sensory innervation of the mystacial pad. I. Innervation of vibrissal follicle-sinus complexes. J Comp Neurol 252: 154–74

Sachdev R, Berg R, Champney G, Kleinfeld D, Ebner F (2003) Unilateral vibrissa contact: changes in amplitude but not timing of rhythmic whisking. Somatosens Mot Res 20: 163–9

Sachdev R, Sato T, Ebner F (2002) Divergent movement of adjacent vibrissae. J Neurophysiol 87: 1440–8

Shoykhet M, Doherty D, Simons D (2000) Coding of deflection velocity and amplitude by vibrissa primary afferent neurons: implications for higher level processing. Somatosens Mot Res 17: 171–80

Silveira LCL, Heywood CA, Cowey A (1987) Contrast sensitivity and visual-acuity of the pigmented rat determined electrophysiologically. Vision Res 27: 1719–31

Solomon JH, Hartmann MJ(2006) Sensing features with robotic vibrissae. Nature 443: 525

Stüttgen MC, Kullmann S, Schwarz C (2008) Responses of rat trigeminal ganglion neurons to longitudinal vibrissa stimulation. J Neurophysiol 100: 1879–84

Stüttgen MC, Ruter J, Schwarz C (2006) Two psychophysical channels of vibrissa deflection in rats align with two neuronal classes of primary afferents. J Neurosci 26: 7933–41

Szwed M, Bagdasarian K, Ahissar E (2003) Encoding of vibrissal active touch. Neuron 40: 621–30

Szwed M, Bagdasarian K, Blumenfeld B, Barak O, Derdikman D, Ahissar E (2006) Responses of trigeminal ganglion neurons to the radial distance of contact during active vibrissal touch. J Neurophysiol 95: 791–802

Temereanca S, Simons DJ (2003) Local field potentials and the encoding of vibrissa deflections by population firing synchrony in thalamic barreloids. J Neurophysiol 89: 2137–45

Towal RB, Hartmann, MJ (2006) Right-left asym-

metries in the whisking behavior of rats anticipate head movements. J Neurosci 26: 8838–46

Towal RB, Hartmann MJZ (2008) Variability in velocity profiles during the free-air whisking behavior of unrestrained rats. J Neurophysiol 100: 740–752

Towal RB, Hartmann MJZ (2010) Principles and Applications of Active Tactile Sensing Strategies in The Rat Vibrissal System. IEEE Sensors Conference. Kona, HI. Nov 01–04, 2010

Towal RB, Quist BW, Gopal V, Solomon JH and Hartmann MJZ (2011) The morphology of the rat vibrissal array: a model for quantifying spatiotemporal patterns of whisker-object contact. PLoS Computational Biology 7:e1001120.

Van der Loos H (1976) Barreloids in mouse somatosensory thalamus. Neurosci Lett 2: 1–6

Vaziri A, Jenks RA, Boloori AR, Stanley GB (2007) Flexible probes for characterizing surface topology: From biology to technology. Exp Mech 47: 417–25

Vincent SB (1913) The tactile hair of the white rat. J Comp Neurol 23: 1–38

Welker WI (1964) Analysis of sniffing of the albino rat. Behav 22: 223–244

Williams C, Kramer E (2010) The advantages of a tapered vibrissa. PLoS One 5: e8806. doi:10 1371/ journal. pone.0 008 806

Wineski L (1983) Movements of the cranial vibrissae in the golden hamster (*Mesocricetus auratus*). J Zool (Lond) 200: 261–80

Wineski LE, Donald MR, Pitts SA (1988) Morphology of the vibrissal motor system in rodents exhibiting different types of exploratory behavior. Amer Zool 28: 77A

Woolsey TA, Welker C, Schwartz RH (1975) Comparative anatomical studies of SMl face cortex with special reference to occurrence of barrels in layer 4. J Comp Neurol 164: 79–94

Zucker E, Welker WI (1969) Coding of somatic sensory input by vibrissae neurons in the rat's trigeminal ganglion. Brain Res 12: 138–56

Touch mechanoreceptors: modeling and simulating the skin and receptors to predict the timing of action potentials

15

Gregory J. Gerling, Daine R. Lesniak, Elmer K. Kim

Contents

Abstract

The sense of touch, and the neural code that underlies it, makes it possible to execute everyday tasks such as picking up a glass or buttoning a shirt. Recreating touch in upper limb prostheses requires an in-depth understanding of how indentation of the skin is transformed to the time of occurrence of neural action potentials by tactile mechanoreceptors, and this understanding may come from modeling and simulation. Typical models transform either sustained indentation to firing rates, or vibration to the timing of action potentials. In contrast to those approaches, the skin-receptor model overviewed here simulates the transfer function of the slowly adapting type I (SA- I) receptor with input of indentation and output of the timing of action potentials. To achieve this transformation, the model couples a finite element model of skin, a sigmoidal transduction function, and a leaky integrate-and-fire neuron model. Subunits of the computational model are then enhanced and combined with a physical sensor to produce action potential discharge that is differentially responsive to the ramp and hold phases of indentation. These are compared to recordings from a mouse SA-I receptor. The overall approach may aid sensor and systems designers building touch prosthetics in an effort to restore perception at the peripheral afferent.

Gregory J. Gerling
University of Virginia, Department of Systems and Information Engineering
151 Engineer's Way, Charlottesville, VA 22904
e-mail: gregory-gerling@virginia.edu

1. Introduction

The sense of touch, and the neural code that underlies it, makes it possible to execute everyday tasks such as picking up a glass or buttoning a shirt. In persons afflicted by diseases (diabetes and AIDS), lack of touch and pain sensation can lead to injuries that result in irreversible tissue damage and chronic pain. Understanding the sense of touch is important for designing human-machine interfaces in medical environments, for example, to aid doctors who operate surgical robots and perform manual examinations. A greater understanding of touch is also required for restoring touch sensation via upper limb prostheses which will interface with the human nervous system.

We have begun to understand touch sensation via both electrophysiological experiments where neural responses to stimulus properties are recorded and psychophysical experiments where perceptual responses to stimulus properties are solicited from human subjects. These two experimental methodologies are in many ways complementary as some questions only apply at one level. The integration of the two can help elucidate underlying physiological and neural mechanisms. However, as neither measuring nor precisely controlling the forces within the skin is currently possible, an understanding of how a stimulus, the skin, and the receptor interact to generate trains of action potentials must be gained through computational modeling. Engineering and physiology intersect here and engineers often have to learn the fine art of the possible in their colleague's domain, and vice versa. The intention of this chapter is to overview the neurophysiology of touch sensation and some models of the skin, receptor, and neural response. In more detail is discussed a model of the slowly adapting type I mechanoreceptor that seeks to predict the sequences of action potentials elicited by mechanical indentation.

1.1 Neurophysiology of touch sensation

Our sense of touch is mediated by arrays of touch-sensitive receptors embedded in skin. Mammalian cutaneous mechanoreceptors are classified by the type of stimulus which elicits their preferred response, and also by their morphology, receptive field characteristics, rate of adaptation, and psychophysical sensation (Johnson et al. 2000; Johnson 2001). In general, Pacinian corpuscles (PC) respond to vibration, rapidly adapting (RA) receptors respond to flutter and moving stimuli, slowly adapting type II (SA-II) receptors respond to stretch, and slowly adapting type I (SA-I) receptors respond to sustained and vibratory indentation.

The slowly adapting type I (SA-I) mechanoreceptor, the focus of this work, facilitates the perception of curvature and edges which enables us to grasp objects such as glasses, buttons, and spoons (Phillips et al. 1981). SA-I receptors are critical for resolving spatial detail and discriminating surface form and orientation. Additionally, SA-I receptors may become functionally more important as we age (Cole et al. 1999) and their transduction mechanism may be closely related to pain reception.

The SA-I transduction end organ is a cluster of Merkel cells, whereby a single SA-I afferent fiber branches to innervate ~5–40 Merkel cells. Merkel cells are located in the basal layer of the epidermis near the epidermal-dermal stiffness border, and are found in both the glabrous skin of fingertips (Fig. 1) and the touch domes of hairy skin (Guinard et al. 1998).

When an object contacts the skin's surface, mechanical forces propagate through the underlying tissue toward the locations of SA-I mechanoreceptors. Each SA-I encodes

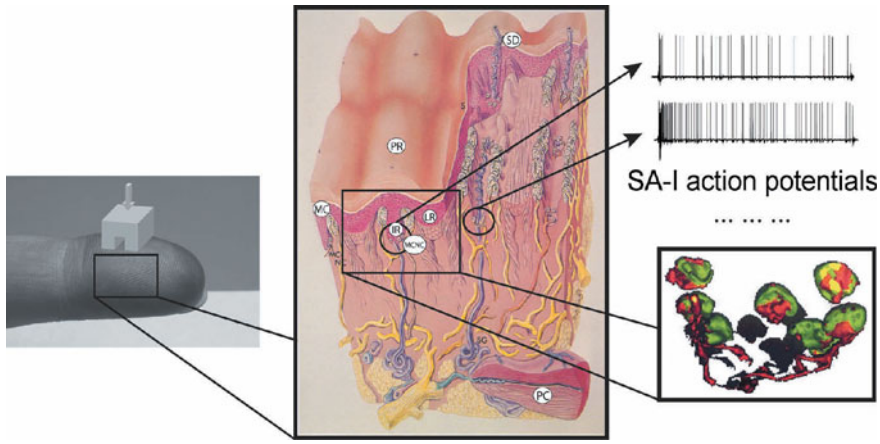


Fig. 1 Basic neurophysiology of touch, showing (left) indentation at the surface of the skin, (center) layers of the skin with the locations of the MCNC (Merkel cell – neurite complex) near the IR (Intermediate Ridge) that lies in opposition to the PR (Papillary Ridge), (upper right) two example trains of action potentials, where a more vigorous response might correspond to an SA-I afferent under a larger magnitude of force, and (lower right) 3D reconstruction of a cluster of Merkel cells using confocal scanning laser microscopy. Abbreviations MC, SD, and PC stand for Meissner corpuscle, sweat duct and Pacinian corpuscle. The insets are adapted from Dillon (1981) and Guinard et al. (1998)

mechanical distortion local to its end organ into trains of neural pulses, or spikes, in a process termed tactile mechanotransduction (Johansson et al. 1979; Mills et al. 1995; Ogawa 1996; Tachibana et al. 2002; Halata et al. 2003), (Fig. 1). Once encoded, trains of action potentials from hundreds of SA-I fibers are sent to the central nervous system (CNS). When decoded by the CNS, the outcome is a three-dimensional neural representation of the deformed skin surface, similar to a retinal image (Gardner et al. 2000; Goodwin et al. 2004). An example spike train recorded from a single SA-I in response to an applied force is shown (Fig. 2).

Electrophysiological and psychophysical experiments in humans or monkeys have been used to determine the limits of absolute and differential spatial acuity mostly using square gratings (Craig 1999; Wheat et al. 2000; Gibson et al. 2002; Goldreich et al. 2003; Sanger 2003; Vega-Bermudez et al. 2004; Gibson et al. 2005; Ingeholm et al. 2006), curvature (Goodwin et al. 1991; Goodwin et al. 1997; Wheat et al. 2001), and small dot stimuli (LaMotte et al. 1986; Srinivasan et al. 1990; Phillips et al. 1992) applied statically. The application of dynamic (vibratory) stimuli also has a rich history (Bolanowski et al. 1984; Whang et al. 1991; Kaczmarek et al. 2000).

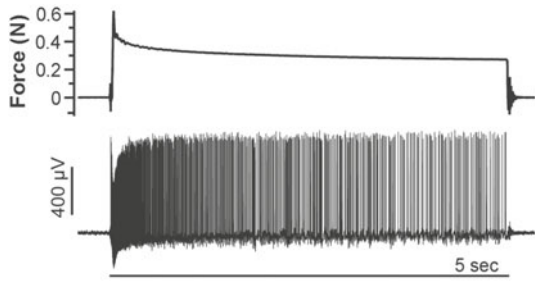


Fig. 2 Example extracellular recording of neural action potentials from an SA-I afferent in response to the force shown applied with a 3 mm diameter cylinder probe. Adapted from Lesniak et al. (2009b)

Apart from *in vivo* experiments, input-output relationships for touch receptors (both slowly and rapidly adapting) have been investigated using isolated skin-nerve preparations in animals. Once a skin flap is separated from the body and placed on a flat substrate, compressive or tensile stimuli are applied to the skin and the response at a single afferent nerve is recorded. Khalsa, Hoffman, and Grigg, in specific, have used this preparation to afford the precise delivery of stimuli. Their specially-built apparatus can stretch and hold the skin in various mechanical states of stress; to investigate, for example, how various bi-axial states of stress, as well as shear stresses, impact the vigorousness of the neural response (Khalsa et al. 1996). Additionally, with the ability to control indentation normal to the skin surface, it was shown that indentation by control of compressive stress leads to more vigorous firing in SA-I receptors than by displacement, force, or strain (Ge et al. 2002).

Although we understand, in general, how skin receptors convert skin deformation into a neural response (Johnson et al. 2000; Johnson 2001) and how responses from a population of receptors are interpreted by the brain (Burton et al. 2000; Gardner et al. 2000), many practical details have yet to be resolved. As a result, we are not able to reproduce how a population of receptors responds to indentation stimuli, positioned randomly on the fingerpad, nor are we remotely close (Goodwin et al. 2004). This is in part because simultaneous mapping of responses from hundreds of receptors is not currently possible and because skin mechanics and receptor transduction are complex while models are primitive. Moreover, we do not entirely understand what features are embedded in the neural response, nor how the CNS interprets the trains of action potentials from the sensory periphery (Johansson et al. 2004).

1.2 Models of touch sensation

Models to predict the neural response to tactile stimuli focus mainly upon either skin mechanics or neural dynamics. Both types of models will be overviewed, in terms of input-output relationships and ability to predict the spatial and/or temporal characteristics of an imposed stimulus.

Models of skin mechanics have investigated how the skin embeds edge features into the neural response. The skin mechanics models indicate SA-Is respond more vigorously to edges because higher magnitudes of stress and strain concentrate in the skin beneath indented edges. Skin mechanics are typically modeled under the assumption that the receptors are stress or strain sensors in an elastic skin matrix using either continuum mechanics (Phillips et al. 1981; Sripati et al. 2006) or finite elements (Srinivasan et al. 1996; Maeno et al. 1998; Dandekar et al. 2003; Wu et al. 2004). Basically, an indenter deforms simulated skin and distributions of stress and strain are calculated in underlying tissue. Various stress and strain measures, such as strain energy density (SED), are then sampled at the sensor location and converted via a scaling function to SA-I firing rates. Predicted SA-I firing rates are typically compared through correlation to *in vivo* firing rates, for like stimuli, recorded from single touch receptors by inserting a microelectrode into a single peripheral nerve fiber at the arm (Johansson et al. 1979). These models have been used to investigate the perception of slip in grasping tasks (Maeno et al. 2003), detection limits for grating, gap, and curved stimuli (Srinivasan et al. 1996; Wu et al. 2005), dynamic contact and force response (Serina et al. 1998; Pawluk et al. 1999), and skin viscoelasticity (Wu et al. 2004).

The major limitation of skin mechanics models as previously employed is that they fit only to firing rate, which may not fully

explain the sequences of action potentials from which the brain derives perceptual information. In particular, recent studies suggest that the timing of individual action potentials must be considered to explain the transmission of perceptual information within behavioral timescales. For example, people can react to touched objects within 100 msec, when the brain has only had time to process the first few spikes (Johansson et al. 2004). The timely arrival of just the first pulse (~40–80 msec after stimulus onset) may tend to fit the behavioral timescales necessary for the recognition of friction and object curvature (~100 msec) (Westling et al. 1984; Jenmalm et al. 1997). Indeed, the timely identification of stimulus onset events has been found to be similarly important for auditory (Heil 2004) and other sensory modalities (VanRullen et al. 2005). Since skin elasticity models only predict average discharge rate, rather than the occurrence of individual action potentials over adaptation time, these models cannot be used to investigate other coding metrics, e. g., first action potential latency (Johansson et al. 2004) or action potential discharge, in early versus later phases of a ramp and hold stimulus. To do so, the existing paradigm of combining skin mechanics models with scaling functions must be extended.

One way to extend skin mechanics models to predict the time of occurrence of action potentials is by modeling SA-I neural dynamics. Generally, neural dynamics are modeled using either the Hodgkin-Huxley model, two-dimensional reductions of the Hodgkin-Huxley model, or leaky integrate and fire models (Gerstner 2002; Izhikevich 2004). SA-I neural dynamics have predominantly been modeled with specialized leaky integrate and fire models (Freeman et al. 1982 a; Freeman et al. 1982 b; Slavik et al. 1995; Leung et al. 2005). These models approximate the SA-I membrane as a resistive capacitive (RC) circuit (similar to other cell membranes) to predict the time of occurrence of action potentials elicited by vibration when the membrane potential is driven to threshold. By defining a discharge threshold, the time of response onset can be predicted. These models, however, tend to not include skin mechanics or other nonnervous aspects of stimulus transformation. This is a major problem because the skin mechanics, receptor end-organ transduction, and neural dynamics are inherently tied together (Darian-Smith 1984; Pubols et al. 1986; Pubols 1988). One other limitation is that the input of these models is limited to vibratory stimuli, at the exclusion of sustained stimuli.

In contrast to the aforementioned efforts in computational modeling, others are attempting to build physical models, where artificial sensors are positioned in elastic substrates, to provide tactile feedback to a user. For example, tactile sensors modeling the skin and ridges have been built for use in robotic applications (Maeno et al. 2003). While able to provide information on stimulus magnitude, typical force sensors lack the capability to produce neural spikes. Instead, force sensors typically report continuous change in voltage over time. It may be beneficial to convert a sensor's voltage response into trains of action potentials if integrated with the nervous system.

In this chapter, both computational modeling and physical modeling efforts are addressed. The first is a computational model of the SA-I afferent that is able to output trains of action potentials in response to skin indentation. The second is a physical model that takes the output from a force sensor and transforms its voltage response into trains of action potentials in a way that is differentially responsive to ramp and hold phases of indentation. These algorithms, combined with artificial sensors, might enable tactile prostheses in upper limb amputees. Such prostheses would provide crucial feedback

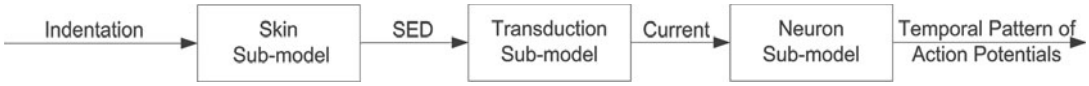


Fig. 3 Block diagram of components of skin-receptor model

for performing everyday tasks in a fashion similar to that of mechanoreceptors in our glabrous skin.

2. A computational model of the SA-I afferent to produce spike-based output

The skin-receptor model described here is a transfer function to convert skin indentation to the timing of action potentials (Lesniak et al. 2009a), thereby bridging the gap between skin mechanics and neural dynamics models. The skin-receptor model is composed of three coupled sub-models: a skin mechanics sub-model, a transduction sub-model, and a neural dynamics sub-model (Fig. 3). Each sub-model performs a sub-transformation towards predicting SA-I spike times elicited by skin indentation.

2.1 Skin mechanics – transduction – neural dynamics model

Skin indentation is transformed into strain energy density (SED) at the receptor by the skin sub-model. The skin is modeled with a two-dimensional finite element model (FEM) that captures the geometry and properties of fingertip skin (Fig. 4) (Gerling et al. 2008). Specifically, stiffness and thickness of the epidermal, dermal, and hypodermal layers are present, as is the undulating interface between the epidermal and dermal layers. The troughs of this undulating interface are

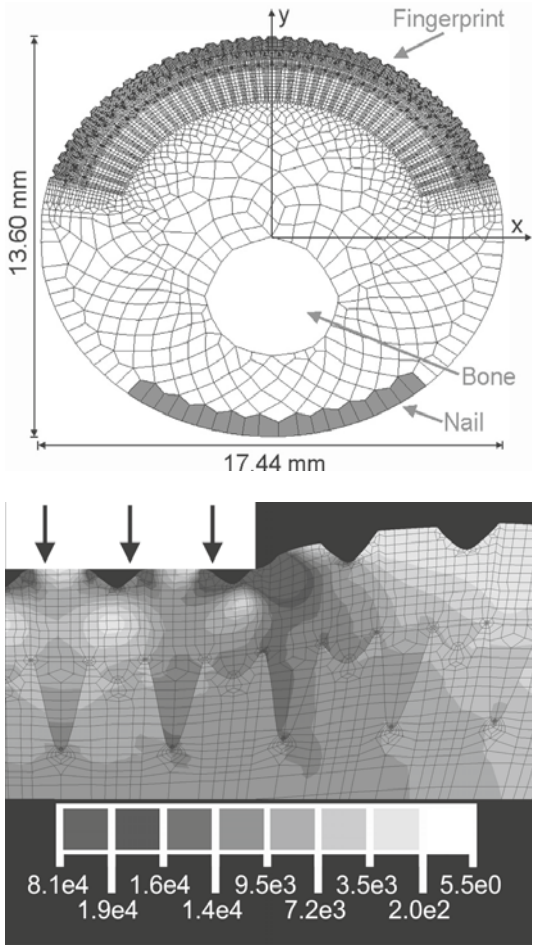


Fig. 4 (upper panel) Geometry and mesh of the ~16 000 element, finite element model which represents a cross-section of the fingerpad and **(lower panel)** deformed mesh, under the load of an indenter displaced to 1.0 mm in the direction normal to the fingerpad surface, showing the propagation of the resultant von Mises stress (units of Pascals) through the tissue layers. Adapted from Gerling et al. (2008)

known as intermediate ridges, and the tips of intermediate ridges are where SA-I receptors are located. Therefore, SED measured at

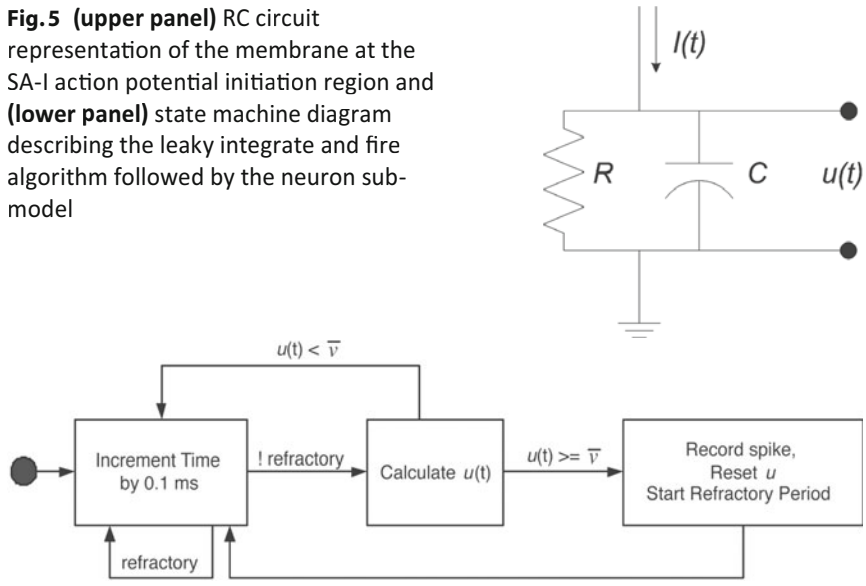


Fig. 5 (upper panel) RC circuit representation of the membrane at the SA-I action potential initiation region and **(lower panel)** state machine diagram describing the leaky integrate and fire algorithm followed by the neuron sub-model

an intermediate ridge tip represents the SED distorting a SA-I receptor, and is the quantity passed to the transduction sub-model.

The transduction sub-model transforms the SED at the SA-I receptor to transmembrane current. Since the receptor is not accessible to electrophysiological recording, the relationship between SED and transmembrane current is unknown. However, sensory cells such as hair cells and pain receptors exhibit stimulus-current curves that are sigmoidal (Holt et al. 2000; Siemens et al. 2006). For this reason, the transduction sub-model utilizes a sigmoidal function (Eq. (1)) where α, β, and γ specify the shape of the function, and where ε is the SED that gives rise to the transmembrane current, I, passed to the neural dynamics sub-model.

$$I(\varepsilon) = \alpha \frac{1}{1 + e^{\gamma(\beta - \varepsilon)}} \tag{1}$$

The neural dynamics sub-model generates action potentials in response to transmembrane current using a leaky integrate and fire model that represents the SA-I membrane as an RC circuit (Fig. 5, upper panel) (Gerstner 2002). The dynamics of this RC circuit (Eq. (2)) define how transmembrane current entering the neuron sub-model generates a membrane potential, where τ is the membrane RC time constant.

$$\tau \frac{\partial u}{\partial t} = -u(t) + RI(t) \tag{2}$$

When the membrane potential is driven to a threshold value, the membrane potential is reset, an absolute refractory period is entered, and a time for the occurrence of an action potential is recorded.

Algorithmically (Fig. 5, lower panel), the neuron sub-model simulates time in 0.1 msec increments. For each time increment, the neuron sub-model is either in a refractory period (refractory) and does nothing, or is not in a refractory period (! refractory) and calculates the membrane potential using Eq. (2). Whenever the membrane potential reaches threshold, an action potential is recorded and the 1.0 msec refractory period is entered.

2.2 Example results

Example results from the model are shown in Figure 6, for indentation depths varying from 1.0 to 2.0 mm for a 3.0 mm bar indenter. In the rightmost plot, note changes of the latency of the first action potential and of the firing rate (or inter-spike interval).

This model is able to predict spike timing, in addition to firing rate. While shown with a sustained stimulus here, the model could be used with a dynamic stimulus or incorporate skin viscoelasticity. In particular, the model is extended in *Section 3* to reproduce the high firing frequency of the SA-I receptor in the stimulus ramp up, as compared to its lower firing frequency when the stimulus is held in place.

3. Implementation of a spiking-sensor model with a force sensor in a silicone substrate to differentially respond to ramp and hold phases of indentation

To extend the computational model into the physical realm, demonstrated here is the integration of the aforementioned transduction and neural dynamics sub-models with an artificial force sensor embedded in a silicone-elastomer substrate (Kim et al. 2010). The intent is to produce a spiking response from a force sensor that typically reports change in voltage. In addition, the previous transduction function is refined to produce action potential discharge that is differentially responsive to ramp and hold phases of indentation (i. e, sensitive to rate of indentation). The SA-I's "spike firing" has two distinct phases in response to ramp and hold indentation: dynamic and static (Johnson 2001). The dynamic phase, i. e., when the stimulus is moving, produces a burst of action potentials with a high discharge rate. The static phase, i. e., when the stimulus probe has stopped moving and is held into the skin, produces a comparably lower discharge rate. We note to the reader that at the onset of stimulus ramp-up and nearing its hold phase, there typically are observed considerable acceleration components. These are not included here in our definition of ramp and hold.

An experiment used the spiking-sensor model to predict times of action potential occurrence for three indentation depths. The objectives were to a) demonstrate the production of a response sensitive to the rate of indentation and b) compare the responses to recordings from isolated skin-nerve preparations. The predicted results

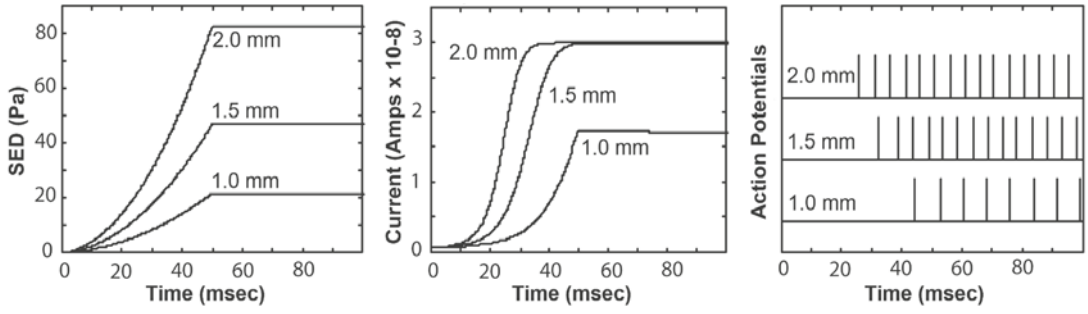


Fig. 6 Transformation of SED to current to action potentials, over time, for 1.0, 1.5, and 2.0 mm deep indentations. Note in the right hand plot, the lack of decline in the action potential rate for the static phase of the stimulus, which is unnatural and does not well represent adaptation in comparison to the high frequency initial discharge that is expected

were compared to neural recordings from a single SA-I fiber in the mouse under similar indentations, using a protocol described elsewhere (Lesniak et al. 2009b; Maricich et al. 2009). To evaluate part b), we analyzed the latency of the first action potential (time interval between onset of stimulus and onset of first action potential) and discharge rate (using inter-spike interval) over both the dynamic ramp-up and static sustained indentation phases.

3.1 Spiking-sensor model

3.1.1 Force sensor – silicone substrate

A single, piezoresistive force sensor was embedded into a silicone-elastomer substrate such that controlled compressive stimuli elicited time-changing voltage output. The commercially available sensor (Flexiforce A201, Tekscan Inc., South Boston, MA) responds to normal force in a 0–4.4 N range over its thin (0.20 mm) and circular (diameter 9.53 mm) pressure sensitive area. When the sensor is integrated into a force-to-voltage circuit, the force applied on the pressure sensitive area reduces the resistance across the sensor so that applied force translates to a change in the measured voltage.

The single sensor was embedded at a laterally central location in the cylindrical-shaped, silicone-elastomer substrate (diameter 30 mm, height 10 mm) and at a depth of 1.0 mm from the surface. The stiffness of silicone-elastomer was set to a 136 kPa elastic modulus, as this value closely matches that reported for the epidermal layers of skin in human cadavers (Maeno et al. 1998).

3.1.2 Spike generation model (transduction and neural dynamics)

The transduction and neural dynamics models were connected to the output of the force sensor, and thereby used as the algorithm underlying the spiking-sensor model. However, the aforementioned transduction model was modified to represent the rate sensitive response of the SA-I neuron. As before, force detected at the sensor is transformed into current, similar to how applied stress/strain at an SA-I afferent is transformed into receptor current across its membrane. However, while the prior model considered only static phase spiking through the use of a sigmoidal function, we here add the ability to separately produce spike bursting in the first 100 msec. Therefore, the transduction function (Eq. (3), (4)) developed here transforms force into current, where $I(t)$ is the transmembrane current, $f(t)$ is the sensor detected force, $df(t)$ is the change in sensor detected force, β is the intercept constant, k_s is the static gain, k_d is the dynamic gain, and Δ is the resolution or step size of time of force detected (10 msec in our case).

$$I(t) = \beta + k_s f(t) + k_d df(t) \tag{3}$$

$$df(t) = \begin{Bmatrix} 0, & \text{for t = onset} \\ \left| \dfrac{f(t) - f(t-\Delta)}{\Delta} \right|, & \text{otherwise} \end{Bmatrix} \tag{4}$$

The static term of the sub-model linearly relates to the force and the dynamic term relates to the magnitude of the first-order force differences. Constant terms β (0), k_d (3 E-03), and k_s (3 E-07) were determined through model fitting, and the dynamic term was found to dominate in the < 500 msec phase while the static term dominates thereafter. Note also that the neural dynamics from Section 2 were used here with parameters of the τ (72 219 msec), C (4.78 E-6 mF), and $\bar{v}$ (41 476 mA) as determined through model fitting.

3.2 Experimental procedure

A set of experiments were conducted with the spiking-sensor model where a sustained displacement of the flat end of a rigid cylindrical rod (20 mm diameter, polyoxymethylene material commonly referred to as Delrin) was delivered at three depths (1.1, 1.3, 1.5 mm) within ~100 msec and held for 5 sec. The indenter tip had a speed of 5.2 (±0.7) mm/sec, no curvature and was centered over the sensor's receptive field. The 20 mm tip size ensured that equal force was delivered over the entire receptive field of the force sensor. A load cell (Honeywell, Model Sensotec 11 subminiature, Columbus, OH) recorded normal force with a 44 N maximum load capacity. The load cell was mounted to a mechanical linear z-stage (Newport, Model ILS100) that was controlled by a motion controller (Newport, Model ESP300) with a displacement resolution of 0.1 µm.

Once the sensor-substrate was preconditioned, the indenter tip was lowered until it was judged to be barely contacting the surface of the sensor-substrate with the visual aid of a camera. Next, the indenter tip was displaced at a rate of ~5.2 mm/sec and provided a sustained indentation for 5 sec. After the 5 sec, the tip was released at the same rate and kept idle for 2 sec. This process was repeated where the order of the indentation depths was randomized. Fig. 7 shows an indentation to a depth of 1.5 mm.

To provide biological comparison data, electrophysiological neural recordings from isolated skin-nerve preparations were conducted on a single mouse, similar to a procedure described elsewhere (Maricich et al. 2009). In addition to recording extracellularly, the force at the probe tip was recorded. The apparatus and experimental runs used the flat end of a cylinder to indent the skin encompassing the entire receptive field. Individual trials consisted of a 5 sec displacement at a speed of ~30 mm/sec. The time

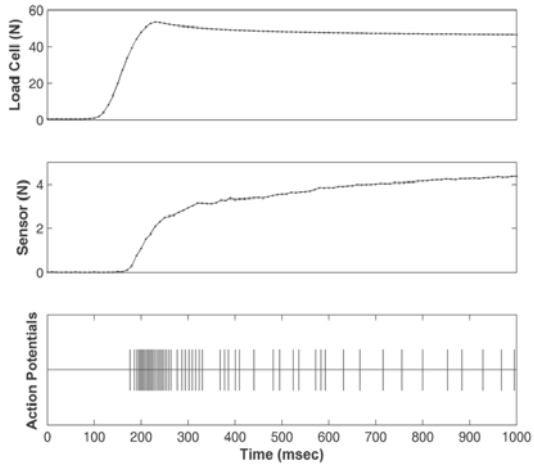


Fig. 7 Plot of data transformation for a single 1.5 mm indentation into the sensor-substrate. The force is recorded at the *load cell* attached to the indenter tip, the force value at the Flexiforce *sensor* in the sensor-substrate, and the production of *action potentials* for that stimulus. Only the first 1.0 sec is shown

to ramp-up was around 50 msec. Because the artificial sensor was larger than the SA-I receptive field, the indenter tip in the artificial sensor preparation was sized at 20 mm, compared to 3 mm in the mouse preparation. This required greater force to traverse into the substrate and therefore slowed the maximum rate of the actuator, to 5.2 mm/sec compared with 30 mm/sec in the mouse preparation.

3.3 Example results

The results of Fig. 8 demonstrate the desired rate sensitivity and the linear increase in firing rate (decrease in inter-spike interval) with increasing indentation depth (Johnson 2001). Over the three displacements 1.1, 1.3, and 1.5 mm, the average static inter-spike interval (ISI) was 69.0, 45.2, 35.1 msec for the spiking-sensor model and 159.9, 67.2, 35.4 msec for the SA-I receptor, while average dynamic ISI was 7.3, 4.2, 3.8 msec and

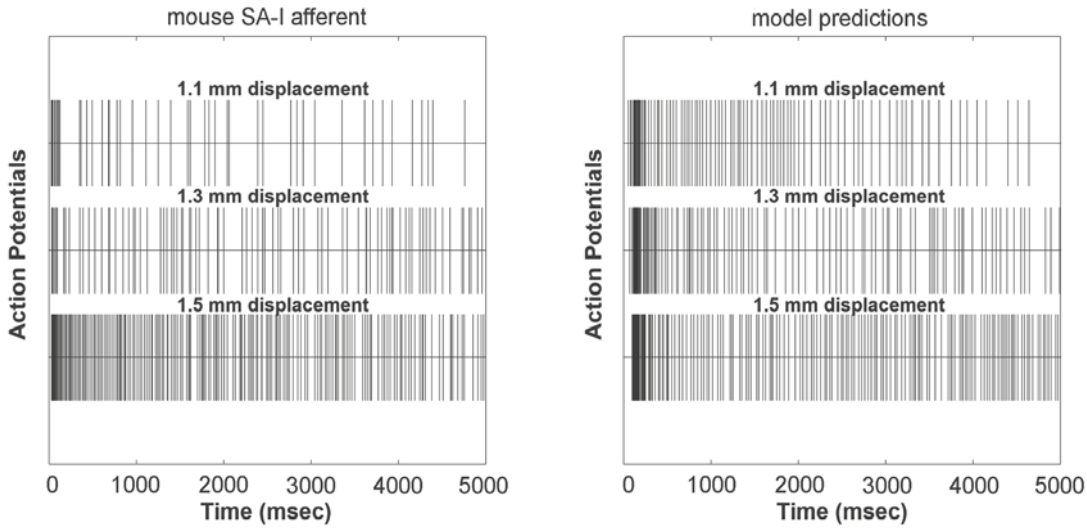


Fig. 8 Trace depicting the timing of action potentials for indentation at three displacement values. Shown (left) is a series of timings of action potentials recorded from one SA-I receptor via an *ex vivo* mouse prep and (right) a series of spike times predicted by using the sensor-substrate and the spiking-sensor model. Note that the average dynamic ISI was calculated as the sum of the ISIs from the occurrence of the first spike to the spike time immediately before the time of max load divided by the number of ISIs within this window, for both the artificial and neural responses. Furthermore, average static ISI was calculated similarly over a 2–5 sec window for both the artificial and neural responses

6.2, 7.1, and 3.8 msec, respectively. Note that average static ISI was calculated over the 2–5 sec interval, well into the stimulus hold and not including the high discharge of action potentials when the stimulus was backed off at 5 sec. The ISI values produced by the artificial system were clearly different between its dynamic and static phases and decreased as indentation depth increased. In addition, the predicted ISI values were within the vicinity of those observed from the mouse SA-I, though lower in the dynamic phase for the artificial system.

However, the predicted first spike latencies were longer (74.3, 73.9, 96.3 msec) in comparison to those observed in the SA-I receptor (26.8, 31.7, 28.8 msec). The predicted later occurrence of the first action potential as compared to the mouse SA-I receptor is due, in part, to the different ramp-up rate and duration, which measured 5.2 mm/sec, ~100 msec in case of the artificial sensor and 30 mm/sec, ~50 msec in the mouse preparation. That said, the mouse SA-I produces an exquisitely rapid response compared to the artificial force sensor's response lag (see lag in Fig. 8 between onset for *Load Cell* and *Sensor*).

Finally, Fig. 8 illustrates the irregular nature of the spike-to-spike elicitations, similar to that of the observed traces and known to accompany SA-I afferents. This effect is produced in the model due to the amplification of the small changes in force, or perhaps to noise in the circuit, originating from the sensor.

Conclusions and future directions

Recreating touch in upper limb prostheses requires a greater understanding of how skin indentation is transformed to spike times by slowly adapting type I (SA-I) mechanoreceptors. Both computational and physical models may aid in generating optimal solutions for tactile

sensor configuration. The examples here produce spike firing for sustained stimuli and well mimic the rate sensitive response of the SA-I mechanoreceptor to a ramp and hold stimulus. Future efforts will need to consider both sustained and vibrational stimulus prediction and the characteristics of the other three types of tactile mechanoreceptors. Another important future direction will be determining how to interface micro-wires to the nervous system so that artificial sensation is connected to the CNS.

Acknowledgments

The project described was supported by grants from the Defense Advanced Research Projects Agency (DARPA Grant Number HR0011-08-1-0072) and the National Library of Medicine (Grant Number T15LM009462). The content is solely the responsibility of the authors and does not necessarily represent the official views of DARPA, the National Library of Medicine, or the National Institutes of Health. The authors would like to acknowledge E. A. Lumpkin and S. A. Wellnitz, Baylor College of Medicine, for their feedback and help in attaining the data published in previous works. Finally, all animal use referenced here was conducted according to the Guide for the Care and Use of Laboratory Animals (NIH), and was approved by the Animal Care and Use Committee of Baylor College of Medicine and the Department of Defense.

References

Bolanowski SJ, Zwislocki JJ (1984) Intensity and frequency-characteristics of pacinian corpuscles. I. Action-potentials. J Neurophysiol 51: 793–811

Burton H, Sinclair RJ (2000) Attending to and remembering tactile stimuli: A review of brain imaging data and single-neuron responses. J Clin Neurophysiol 17: 575–591

Cole KJ, Rotella DL, Harper JG (1999) Mechanisms for age-related changes of fingertip forces during precision gripping and lifting in adults. J Neurosci 19: 3238–3247

Craig JC (1999) Grating orientation as a measure of tactile spatial acuity. Somatosens Motor Res 16: 197–206

Dandekar K, Raju BI, Srinivasan MA (2003) 3-D finite-element models of human and monkey fingertips to investigate the mechanics of tactile sense. J Biomech Eng 125: 682–691

Darian-Smith I (1984) The sense of touch: Performance and peripheral neural processes. In: Brookhart JM, Mountcastle VB, Darian-Smith I (eds) Handbook of physiology: The nervous system III. Am Physiol Soc, Bethesda, Maryland, pp 739–788

Dillon AL (1981) Evaluation of sensibility and re-education of sensation in the hand. The Williams & Wilkins Company. Baltimore, Maryland

Freeman AW, Johnson KO (1982a) Cutaneous mechanoreceptors in macaque monkey: Temporal discharge patterns evoked by vibration, and a receptor model. J Physiol 323: 21–41

Freeman AW, Johnson KO (1982b) A model accounting for effects of vibratory amplitude on responses of cutaneous mechanoreceptors in macaque monkey. J Physiol 323: 43–64

Gardner EP, Kandel ER (2000) Touch. In: Kandel ER, Schwartz JH, Jessell TM (eds) Principles of neural science. McGraw-Hill, New York, pp 451–471

Ge WQ, Khalsa PS (2002) Encoding of compressive stress during indentation by slowly adapting type I mechanoreceptors in rat hairy skin. J Neurophysiol 87: 1686–1693

Gerling GJ, Thomas GW (2008) Fingerprint lines may not directly affect SA-I mechanoreceptor response. Somatosens Motor Res 25: 61–76

Gerstner K (2002) Spiking neuron models. Cambridge University Press, Cambridge, UK

Gibson GO, Craig JC (2002) Relative roles of spatial and intensive cues in the discrimination of spatial tactile stimuli. Percept Psychophys 64: 1095–1107

Gibson GO, Craig JC (2005) Tactile spatial sensitivity and anisotropy. Percept Psychophys 67: 1061–1079

Goldreich D, Kanics IM (2003) Tactile acuity is enhanced in blindness. J Neurosci 23: 3439–3445

Goodwin AW, John KT, Marceglia AH (1991) Tactile discrimination of curvature by humans using only cutaneous information from the fingerpads. Exp Brain Res 86: 663–672

Goodwin AW, Macefield VG, Bisley JW (1997) Encoding of object curvature by tactile afferents from human fingers. J Neurophysiol 78: 2881–2888

Goodwin AW, Wheat HE (2004) Sensory signals in neural populations underlying tactile perception and manipulation. Annu Rev Neurosci 27: 53–77

Guinard D, Usson Y, Guillermet C, Saxod R (1998) Merkel complexes of human digital skin: Three-dimensional imaging with confocal laser microscopy and double immunofluorescence. J Comp Neurol 398: 98–104

Halata Z, Grim M, Bauman KI (2003) Friedrich Sigmund Merkel and his "Merkel cell", morphology, development, and physiology: Review and new results. Anat Rec A Discov Mol Cell Evol Biol 271A: 225–239

Heil P (2004) First-spike latency of auditory neurons revisited. Curr Opin Neurobiol 14: 461–467

Holt JR, Corey DP (2000) Two mechanisms for transducer adaptation in vertebrate hair cells. Proc Natl Acad Sci U S A 97: 11730–11735

Ingeholm JE, Dold GR, Pfeffer LE, Goldstein SR, Johnson KO, Ide D, Van Boven RW (2006) The helix: A multi-modal tactile stimulator for human functional neuroimaging. Ann Neurol 60: 638–638

Izhikevich EM (2004) Which model to use for cortical spiking neurons? IEEE Trans Neural Netw 15: 1063–1070

Jenmalm P, Johansson RS (1997) Visual and somatosensory information about object shape control manipulative fingertip forces. J Neurosci 17: 4486–4499

Johansson RS, Birznieks I (2004) First spikes in ensembles of human tactile afferents code complex spatial fingertip events. Nat Neurosci 7: 170–177

Johansson RS, Vallbo AB (1979) Tactile sensibility in the human hand: Relative and absolute densities of 4 types of mechanoreceptive units in glabrous skin. J Physiol 286: 283–300

Johnson KO (2001) The roles and functions of cutaneous mechanoreceptors. Curr Opin Neurobiol 11: 455–461

Johnson KO, Yoshioka T, Vega-Bermudez F (2000) Tactile functions of mechanoreceptive afferents innervating the hand. J Clin Neurophysiol 17: 539–558

Kaczmarek KA, Tyler ME, Brisben AJ, Johnson KO (2000) The afferent neural response to electrotactile stimuli: Preliminary results. IEEE Trans Rehabil Eng 8: 268–270

Khalsa PS, Hoffman AH, Grigg P (1996) Mechanical states encoded by stretch-sensitive neurons in feline joint capsule. J Neurophysiol 76: 175–187

Kim EK, Gerling GJ, Wellnitz SA, Lumpkin EA (2010) Using force sensors and neural models to encode tactile stimuli as spike-based responses. In: Proceedings of the IEEE Haptic Interfaces for Virtual Environment and Teleoperator Systems Conference, Waltham, Massachusetts, pp 195–198

LaMotte RH, Whitehouse J (1986) Tactile detection of a dot on a smooth surface: Peripheral neural events. J Neurophysiol 56: 1109–1128

Lesniak DR, Gerling GJ (2009a) Predicting SA-I mechanoreceptor spike times with a skin-neuron model. Math Biosci 220: 15–23

Lesniak DR, Gerling GJ, Wellnitz SA, Lumpkin EA (2009b) Statistical analysis and modeling of variance in the SA-I mechanoreceptor response to sustained indentation. In: Proc IEEE Engineering in Medicine and Biology Conf, Minneapolis, Minnesota, pp 6814–6817

Leung YY, Bensmaia SJ, Hsiao SS, Johnson KO (2005) Time-course of vibratory adaptation and recovery in cutaneous mechanoreceptive afferents. J Neurophysiol 94: 3037–3045

Maeno T, Kawamura T (2003) Geometry design of an elastic finger-shaped sensor for estimating friction coefficient by pressing an object. In: Proc IEEE International Conf on Robotics and Automation, Taipei, Taiwan, pp 1533–1538

Maeno T, Kobayashi K, Yamazaki N (1998) Relationship between the structure of human finger tissue and the location of tactile receptors. JSME Int J, Ser C 41: 94–100

Maricich SM, Wellnitz SA, Nelson AM, Lesniak DR, Gerling GJ, Lumpkin EA, Zoghbi HY (2009) Merkel cells are essential for light-touch responses. Science 324: 1580–1582

Mills LR, Diamond J (1995) Merkel cells are not the mechanosensory transducers in the touch dome of the rat. J Neurocytol 24: 117–34

Ogawa H (1996) The merkel cell as a possible mechanoreceptor cell. Prog Neurobiol 49: 317–334

Pawluk DTV, Howe RD (1999) Dynamic lumped element response of the human fingerpad. J Biomech Eng 121: 178–183

Phillips JR, Johansson RS, Johnson KO (1992) Responses of human mechanoreceptive afferents to embossed dot arrays scanned across fingerpad skin. J Neurosci 12: 827–839

Phillips JR, Johnson KO (1981) Tactile spatial resolution. III. A continuum mechanics model of skin predicting mechanoreceptor responses to bars, edges, and gratings. J Neurphysiol 46: 1204–1225

Pubols BH (1988) Spread of skin deformation and mechanoreceptor discharge. In: Hamann W,

Iggo A (eds) Progress in brain research. Elsevier Science Publishers, B. V. (Biomed Div), New York, pp 263–270

Pubols BH, Benkich ME (1986) Relations between stimulus force, skin displacement, and discharge characteristics of slowly adapting type I cutaneous mechanoreceptors in glabrous skin of squirrel and monkey hand. Somatosens Res 4: 111–125

Sanger TD (2003) Neural population codes. Curr Opin Neurobiol 13: 238–249

Serina ER, Mockensturm E, Mote CD, Rempel D (1998) A structural model of the forced compression of the fingertip pulp. J Biomech 31: 639–646

Siemens J, Zhou S, Piskorowski R, Nikai T, Lumpkin EA, Basbaum AI, King D, Julius D (2006) Spider toxins activate the capsaicin receptor to produce inflammatory pain. Nature 444: 208–212

Slavik P, Bell J (1995) A mechanoreceptor model for rapidly and slowly adapting afferents subjected to periodic vibratory stimuli. Math Biosci 130: 1–23

Srinivasan MA, Dandekar K (1996) An investigation of the mechanics of tactile sense using two-dimensional models of the primate fingertip. J Biomech Eng 118: 48–55

Srinivasan MA, Whitehouse JM, LaMotte RH (1990) Tactile detection of slip: surface microgeometry and peripheral neural codes. J Neurophysiol 63: 1323–1332

Sripati AP, Bensmaia SJ, Johnson KO (2006) A continuum mechanical model of mechanoreceptive afferent responses to indented spatial patterns. J Neurophysiol 95: 3852–3864

Tachibana T, Nawa T (2002) Recent progress in studies on Merkel cell biology. Anat Sci Int 77: 26–33

VanRullen R, Guyonneau R, Thorpe SJ (2005) Spike times make sense. Trends Neurosci 28: 1–4

Vega-Bermudez F, Johnson KO (2004) Fingertip skin conformance accounts, in part, for differences in tactile spatial acuity in young subjects, but not for the decline in spatial acuity with aging. Percept Psychophys 66: 60–67

Westling G, Johansson RS (1984) Factors influencing the force control during precision grip. Exp Brain Res 53: 277–284

Whang KC, Burton H, Shulman GL (1991) Selective attention in vibrotactile tasks: Detecting the presence and absence of amplitude change. Percept Psychophys 50: 157–165

Wheat H, Goodwin A (2001) Tactile discrimination of edge shape: Limits on spatial resolution imposed by parameters of the peripheral neural population. J Neurosci 21: 7751–7763

Wheat HE, Goodwin AW (2000) Tactile discrimination of gaps by slowly adapting afferents: Effects of population parameters and anisotropy in the fingerpad. J Neurophysiol 84: 1430–1444

Wu JZ, Dong RG (2005) Analysis of the contact interactions between fingertips and objects with different surface curvatures. Proc IME H J Eng Med 219: 89–103

Wu JZ, Dong RG, Rakheja S, Schopper AW, Smutz WP (2004) A structural fingertip model for simulating of the biomechanics of tactile sensation. Med Eng Phys 26: 165–175

Assessing the mechanical response of groups of arthropod filiform flow sensors

16

Bree Cummins, Tomas Gedeon

Contents

Abstract

Many arthropod filiform flow sensors contain hundreds of sensilla that respond to complex natural airflow stimuli. The combination of complex morphology and richness of natural stimuli present a significant challenge for modeling these sensory apparatuses. We survey advances in our understanding over the last five years of the interactions between filiform hairs immersed in airflow. We then discuss a general characterization of filiform hair response based on our work modeling the cricket cercal system, which utilizes a newly developed unsteady Stokes model to investigate the response of a group of hairs to temporally complex signals.

Bree Cummins
Tulane University
Center for Computational Science
Stanley Thomas Hall, Room 401
New Orleans, LA 70118, USA
e-mail: bcummins@tulane.edu

1. Arthropod air velocity sensors

Organisms have developed different types of sensory systems to collect information about the fluid environment that surrounds them. For example, the lateral line in fish senses spatial pressure gradients in water. Human ears transduce air pressure waves into movements of the cochlear fluid and basilar membrane, which in turn physically deflect the stereovilli of hair cells that are coupled to neurons. In many arthropods, there are temporal pressure gradient sensors analogous to ears, but there are also direct air velocity sensors that mechanically transduce air motion into electrical signals in the organism's nervous system. These direct air velocity sensors consist of very thin hairs, called filiform hairs, which deflect in response to air flow and are coupled to sensory neurons at the base of the hair shaft. Arrays of filiform hairs permit arthropods to evaluate information about the direction and dynamic properties of the low-magni-

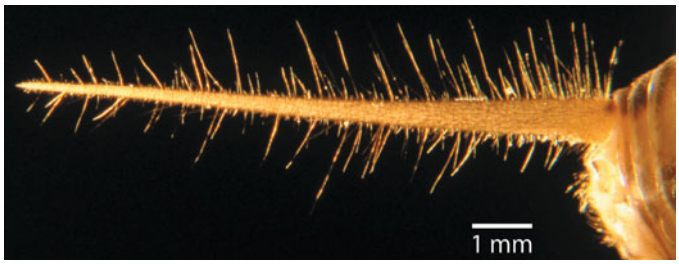


Fig. 1 A single cercus of the cricket *Acheta domesticus* with an approximate 1 mm length scale. Note the dense patch of filiform hairs at the proximal end of the cercus

tude, low-frequency air signals created by predators, prey, or mates. A great deal of effort over the past 30 years has been directed at studying the anatomy, physiological development, and biomechanical and neural response properties of these airflow sensors.

Both the morphology and the arrangement of the mechanosensory hairs vary across arthropod taxa. The hairs of the wandering spider *Cupiennius salei* are elaborately feathered structures that are smoothly curved near the distal end of the hair and are often configured in linear arrays along the spider's legs (Barth et al. 1993). In orb-weaving spiders, the hairs (called trichobothria in spiders) are sparser than in wandering spiders and often do not have the intricate feathering (Barth et al. 1993). Insects, such as the crickets *Acheta domesticus*, *Gryllus bimaculatus*, and *Nemobius sylvestris* and the cockroach *Periplaneta americana*, all have smooth hairs shaped like elongated paraboloids (Kumagai et al. 1998 b) that are distributed over two shallowly-tapering appendages called cerci at the distal end of the abdomen. The insect sensory system comprising the cerci and its population of sensilla, including the filiform hairs, is called the cercal system. The cricket cercal system is larger in size and populated with more filiform hairs than that of many other insects, and includes a prominent and characteristic dense patch of hairs near the proximal end of each cercus (see Fig. 1). Spiders and crickets have been used extensively as model organisms with respect to the biomechanics of filiform hairs, and research on them forms the basis of this discussion.

Despite the morphological differences in filiform hairs seen between taxa, there are functional similarities. The hair sizes are similar, with each species exhibiting a wide range of hair lengths that falls within the confines of about 50 µm to 2 mm. The associated basal diameters range within the bounds of about 1 to 15 µm. The hair lengths act to fractionate frequency responses, with longer hairs responding more vigorously to lower frequencies and shorter hairs to higher frequencies (Barth et al. 1993, Osborne 1996, Kumagai et al. 1998 a). Each of the hairs is seated in a viscoelastic socket that defines the pivot point of the hair and restricts the hair motion. The suspension structure of the hair shaft determines how quickly the hairs relax back into their resting positions after a displacement and how strongly the hairs resist incident airflow. A cuticular structure called the socket cup rises slightly above the surface of the arthropod appendage and limits the maximum possible hair excursion. Additionally, the morphological and mechanical asymmetries of the socket may determine the preferred direction of motion for each hair (Gnatzy and Tautz 1980). The trichobothria of the spider *Cupiennius salei* are able to move much farther (25°–35°) than cricket filiform hairs (5°) before contacting the socket cup (Barth et al. 1993). Trichobothria also exhibit either a mechanical directionality parallel or perpendicular to the long leg axis or, in many cases, a nearly isotropic mechanical response with respect

to the direction of the stimulus airflow (Barth et al. 1993). Cricket hairs always show a strong mechanical directionality (Gnatzy and Tautz 1980), with the most numerous alignments either parallel or perpendicular to the cercus. The filiform hairs of all species respond linearly to sinusoidal airflow that is sufficiently gentle to avoid hair contact with the socket cup.

Slow flows around structures as small as filiform hairs are characterized by fluid motion that is affected more by the internal friction of the fluid than by the inertia. This type of flow is commonly described as a low Reynolds number fluid flow, where the Reynolds number, Re = $U\,L/v$, is the ratio of inertial to viscous forces within the fluid flow. The quantities defining Re are characteristic scales that describe the flow field: U is a characteristic relative velocity, L is a characteristic length scale, and v is the kinematic viscosity of the fluid. For an airflow with a peak velocity of 1 cm/s around a hair with a 5 µm diameter, the Reynolds number is 0.003, assuming that the kinematic viscosity of air is 1.568×10^{-5} m^2/s. Provided that U is small enough to ensure that Re << 1 and that the filiform hairs do not contact the socket cup during their motion, the hairs are reasonably modeled as linear oscillators that move due to drag forces imposed by low Reynolds number fluid flow. We briefly introduce the methodology for calculating these drag forces in the next section, but see Humphrey et al. (1993) and Humphrey and Barth (2008) for in-depth discussions.

Because filiform hairs are so slender and are immersed in slow fluid flows, closely-packed hairs can potentially exhibit viscosity-mediated coupling: interactions between hairs that occur due to the contact between fluid boundary layers that form around the hairs. There has been a long-standing discussion in the literature about the occurrence of significant levels of viscosity-mediated coupling. In many situations, such as in the sparse hair distributions of spiders (Bathellier et al. 2005) or in the low density region near the distal end of cricket cerci, the fluid-mediated coupling between hairs appears to be negligible, so that the hairs may be modeled as independent structures. However, there is evidence that coupling could play a significant role in other hair arrangements, particularly in the dense patch of hairs at the proximal end of cricket cerci (Osborne 1996, Kumagai et al. 1998a). The purpose of this work is to provide a brief review of the efforts to model the response of a group of filiform hairs to both simple and complex stimuli. Additionally, we discuss a method for modeling the response of a group of hairs to temporally complex signals.

2. Modeling individual filiform hairs

Models of group hair mechanical response are elaborations on models of individual hair responses. Tautz (1977) was the first person to recognize that the linearity of filiform hair response and their relative rigidity allows them to be approximated as forced, damped harmonic oscillators in angular displacement,

$$I\ddot{\theta}(t) + R\dot{\theta}(t) + S\theta(t) = \tau(t) \qquad (1)$$

where I is the moment of inertia of a hair, R is the torsional resistance due to the mechanical suspension structures within the viscoelastic socket, S is a torsional restoring constant also due to the socket suspension structures, and $\tau(t)$ is the driving torque acting on the hair. The dots above the angular displacement $\theta(t)$ denote time derivatives. Tautz used a small angle approximation to rewrite Eq. (1) in terms of hair tip displace-

ment, and then analyzed the model assuming a sinusoidal force of unspecified amplitude on the right-hand side. He did not attempt to find the precise form of this force.

Researchers who followed Tautz (Fletcher 1978, Shimozawa and Kanou 1984b, Humphrey et al. 1993, Shimozawa et al. 1998) calculated the explicit forcing on a hair based on models of the incident airflow. The driving torque on the hair, τ(t), is the drag force distribution along the hair times the lever arm integrated over the length of the hair:

$$\tau(t) = \int_0^L f(s,t)s \, ds \tag{2}$$

The particular form of the drag force distribution depends on the approximation used to model airflow. The model that has widespread acceptance today was first introduced by Humphrey et al. (1993), who acknowledged important contributions from the earlier modeling attempts of Fletcher (1978) and Shimozawa and Kanou (1984b). Humphrey et al.'s method relies on the combination of a substrate boundary layer and Stokes' mechanical impedance for a cylinder (Stokes 1851).

Boundary layers are smooth velocity gradients that form between two fluid flows that are prescribed to have different velocities. In the case of an arthropod appendage such as a leg or a cercus, air sticks to the appendage surface and matches its velocity. This is known as a no-slip boundary condition. Because of the no-slip boundary condition, the velocity must smoothly vary from zero at the surface of the appendage to the full magnitude of the bulk flow far from the surface. The boundary layer consists of the part of the velocity gradient that starts at the substrate and ends at the point above the surface where the fluid is moving at 99 % of the free stream velocity. Stokes (1851) showed that the height of the boundary layer that forms over an infinite flat plate varies as $1/\sqrt{f}$ with f denoting the frequency of the oscillating plate. Using a change in variables, this can be shown to be the same case as an infinite plate immersed in oscillating flow of frequency *f*. A similar profile is obtained when considering axial flow over an infinite cylinder (Humphrey et al. 1993). Because of the $1/\sqrt{f}$ dependence, higher frequency flows have smaller boundary layers. And higher magnitude (faster) fluid flows corresponding to higher Reynolds numbers have smaller boundary layers. The heights of the boundary layers at biologically relevant frequencies (1–500 Hz) and velocities (0.1 mm/s to 30 cm/s) are on order the lengths of filiform hairs (Shimozawa and Kanou 1984b, Barth et al. 1993, Humphrey et al. 1993), making the boundary layer velocity profile critical to the calculation of drag forces on the hairs.

The drag force at each point along the hair is calculated using a combination of Stokes' mechanical impedance for a cylinder (Stokes 1851) and the boundary layer profile. Stokes calculated the velocity and pressure fields (u and p, respectively) around a narrow cylinder by explicitly solving a linearization of the Navier-Stokes equations now known as the unsteady Stokes equations:

$$\rho \frac{\partial u}{\partial t} = \mu \Delta u - \nabla p, \quad \nabla \cdot u = 0 \tag{3}$$

In Eq. (3), ρ and μ are the density and dynamic viscosity of the air, respectively. This linearization of the general Navier-Stokes model of fluid flow is valid under the assumption of a low Reynolds number, which holds for slow flows around filiform hairs. The drag force per unit length of a cylinder is the integral of the pressure and shear stress fields around the circumference of the hair, which can be rewritten in terms of the relative velocity between the cylinder and the fluid (Stokes 1851). The relative velocity at a point along the length of the hair is taken to be the difference between the hair's local linear velocity and the local boundary layer

profile (Humphrey et al. 1993). The result is a complex conversion factor between velocity and force at each angular frequency ω:

$$f(\omega) = Z(\omega)u(\omega) \quad (4)$$

where $Z(\omega)$ is called the mechanical impedance of the cylinder-air system (Shimozawa et al. 1998).

This method has been used to model individual hair responses (Humphrey et al. 1993, Osborne 1996, Shimozawa et al. 1998), as well as group hair responses in which the members of a hair population are assumed to move independently (Dangles et al. 2005, Dangles et al. 2006b, Magal et al. 2006).

3. Modeling multiple filiform hairs

Fluid-mediated coupling between neighboring sensilla is negligible when filiform hairs are sparsely distributed. Under this condition, the filtering effect of an array may be approximated as the sum of independently interacting hairs. Applying this assumption to the entire cricket cercus, Magal et al. (2006) developed a model of total cercal canopy response for the wood cricket *Nemobius sylvestris*. They modeled each hair based on the work of Humphrey et al. (1993) and Shimozawa et al. (1998) as an independent linear oscillator in the boundary layer that forms over a flat substrate. They completely characterized the number of hairs and their lengths over *N. sylvestris* cerci and used hair directionality data for *Acheta domesticus* (Landolfa and Jacobs 1995) to estimate the number of hairs sensitive to axial flow from the rear of the cricket. These hairs were assumed to cause an afferent response only at threshold velocities that depend on hair length and stimulus frequency, based on results obtained for G*ryllus bimaculatus* (Shimozawa and Kanou 1984a). Magal et al. summed the maximal hair deflection over all hairs that exceeded their velocity thresholds and multiplied the result by the proportion of hairs with appropriate directions of motion. They assume that the afferent impulse discharge rate is proportional to maximum hair deflection in order to use this calculation as an approximation of the total neural response of the cercal canopy. This response measure has been used by Dangles et al. (2005, 2006b) to compare the cercal canopy response of different *N. sylvestris* instars and of wild *N. sylvestris* populations in distinct environments that are exposed to different types of predation pressure.

The studies by Dangles et al. (2005, 2006b) comparing cercal functionality under different circumstances exemplify the types of studies that are needed in order to understand the evolutionary history, developmental constraints, and phenotypic plasticity of the cercal system. However, groups of hairs may have an effect on the fluid velocity profile other than what is predicted by simply summing the effects of independent hairs. Boundary layers form not only around the arthropod appendage, but also around the filiform hairs themselves. These hair boundary layers can be relatively large with respect to hair diameter in slow flows, and can propagate the effects of hair motion from one near neighbor to another. This fluid-mediated coupling can either cause damping in the motion of near neighbors (Bathellier et al. 2005, 2010; Cummins et al. 2007, Cummins and Gedeon 2007) or it can cause amplified motion, also called synergistic coupling (Cheer and Koehl 1987, Cummins 2009), depending on the orientation and frequency of the airflow.

There is empirical evidence (Osborne 1996, Kumagai et al. 1998a) and theoretical analysis (Bathellier et al. 2005, Cummins et

al. 2007, Cummins and Gedeon 2007, Cummins 2009) that suggests hair coupling does occur on cricket cerci, at least in the densely packed region near the cricket body. Since the afferent response in the model of Magal et al. (2006) is proportional to maximal hair excursion, and hair coupling manifests as changes in that hair excursion, significant coupling may shift total canopy response properties from those estimated under an assumption of independence. It thus makes sense to try and quantify the amount of coupling that occurs in a given hair arrangement on the cerci.

Several studies have attempted to either quantify or qualitatively describe the coupling between thin, closely spaced biological hair arrays. Cheer and Koehl (1987) examined the functional shift that occurs in the bristled feeding appendages of tiny crustaceans called copepods. The bristles, or setae, are linear arrays of roughly cylindrical shapes that are actively moved through the water for the purpose of gathering food particles. There was a long-standing debate about whether or not the setae acted as filters that sieved food out of the water or as scoops that pushed food toward the animal. Cheer and Koehl (1987) developed an analytical model of two cylinders with varied spacing in steady low Reynolds number flow perpendicular to the long axis of the cylinders. They found that decreasing the spacing between the cylinders causes synergistic coupling after they are sufficiently close together. So the cylinder arrays act more like a paddle than a rake at a small spacing and low Reynolds number, and change functionality as either of these parameters increases. The paddle-like activity occurs because in low Reynolds number flow, the boundary layers around each cylinder overlap at close distances and create a pocket of slower moving fluid between them. Cheer and Koehl (1987) suggest that the amount of coupling between the cylinders depends on the Reynolds number of the fluid-structure interaction, the ratio of hair diameter to spacing between the cylinders, and the width of the cylinder array.

Humphrey et al. (1993) originally proposed that coupling between hairs is negligible if the nondimensional spacing between hairs, a/d, is larger than the smaller of the ratio of hair length to hair diameter and the inverse of the Reynolds number, $a/d > \min\{L/d, \mathrm{Re}^{-1}\}$. However, this criterion does not depend on either air frequency or the hair mechanical parameters (I, R, and S), both of which may impact the width of the boundary layer that forms around the hair. Bathellier et al. (2005, 2010) revisited the issue of coupling and developed an intricate model capable of calculating the coupling between hairs arranged linearly, as well as performing some elegant experiments with two isolated trichobothria of the spider *Cupiennius salei*. This model depends on the distance between hairs, the diameter and mechanical properties of the *perturbed* hair, the angular velocity of the *perturbing* hair, and the magnitude and frequency of the free stream airflow. The basis of this model is the analytical calculation of Stokes' mechanical impedance for two cylinders, instead of only one. It differs from the earlier Cheer and Koehl (1987) model in that sinusoidal flows, rather than steady flows, are considered. Using this model, Bathellier et al. (2005, 2010) find that coupling between trichobothria of similar lengths appears to be suppressed at the maximal response of the hairs and that larger coupling is seen between hairs of greatly varying lengths, provided that the hairs are sufficiently close. Using an extension of the model to multiple hairs arranged in a line, they further predict that significant coupling could be seen in short cricket hairs located within the dense hair patch on a cricket cercus.

Bathellier et al. (2005, 2010) present a very interesting approach, but they do not

consider the boundary layer caused by the presence of an arthropod appendage. This substrate boundary layer will affect the formation of the boundary layer around a hair. Since the hair boundary layer is wider in slower flow, we suspect that the method is an underestimate of the coupling between linearly arranged hairs in fluid flowing parallel to the insect appendage, since this causes stagnating flow around the hairs. In flow perpendicular to the appendage, the hairs are exposed to faster flow (Humphrey et al. 1993), and so the coupling may be overestimated in this case. The disadvantage of including a substrate boundary layer in the derivation is that the coupling between linearly arranged hairs can no longer be expressed in an easily evaluated closed form. Despite this issue, the coupling estimate supplied by Bathellier et al. (2005, 2010) is the current best analytical criterion for deciding if fluid-mediated hair interaction is negligible in oscillatory flow.

These efforts to model coupling between hairs have been limited to highly constrained arrangements of hairs. On cricket cerci, where hairs appear in dense aggregations exhibiting various directions of motion, it is probable that these methods do not accurately capture the extent of viscosity-mediated coupling. In order to model arbitrary hair arrangements, Cummins et al. (2007), Cummins and Gedeon (2007), Heys et al. (2008), and Cummins (2009) abandoned attempts to find an analytical solution to the mechanical impedance for multiple hairs and instead presented various methods for predicting hair motion numerically.

Heys et al. (2008) model the system in a penalty immersed boundary framework (Kim and Peskin 2007) that does not require simplifications to the Navier-Stokes equations or to the cercal geometry. Briefly, an immersed boundary method is a numerical approach that simulates the interaction of purely elastic structures with a surrounding fluid domain. The penalty immersed boundary method extends this approach to account for inertial effects by modeling distinct inertial and elastic structures that are coupled with strong springs. This numerical method has substantial benefits; any flow regime may be modeled over any surface, including a realistically shaped appendage, an arbitrarily arranged group of non-rigid filiform hairs, and parts of the arthropod thorax or abdomen. Any coupling between the hairs will be automatically accounted for, through the calculation of the velocity field that arises around the irregular hairy surface. Using this method, Heys et al. (2008) find that two identical cricket hairs with axially aligned planes of motion seem to act as a “super” hair as they get closer together, showing a maximum deflection half that of an isolated hair, and a phase shift that differs by 30 % from that of a single hair.

Despite the advantages of this approach, it has the significant disadvantage that it is computationally expensive, requiring a lot of computing time and a coarse discretization scheme. Our goal in seeking an alternative numerical method has been to decrease the computation time in order to evaluate a large number of hair arrangements at a large number of frequencies. In Cummins et al. (2007) and Cummins and Gedeon (2007), we utilize the method of regularized Stokeslets (Cortez 2001) to numerically solve N linear oscillator equations in a boundary layer flow that is perturbed by the presence of filiform hairs. The forces on the hairs due to the perturbation velocity field are calculated within each time step of an ordinary differential equation (ODE) solver using a steady Stokes approximation of the airflow. This method shows significant coupling between hairs, but because of the steady Stokes approximation, the predictions from this model are valid only at distances less than $\sqrt{\nu/\omega}$, where ν is the kinematic viscosity of air and ω is the

angular frequency of the driving airflow.

To overcome the spatial limitation inherent in the steady model, we developed another implementation based on the method of regularized unsteady Stokeslets (Cortez et al. 2010, Cummins 2009). This method simultaneously solves *N* linear oscillator equations and the unsteady Stokes equations with velocity matching conditions along the hairs, presuming the hairs to be immersed in an analytically calculated boundary layer flow. As in our first method, we model the effect of hairs as a perturbation velocity field that contributes to the forcing distribution along each hair, but an ODE solver is not required. Although this method does not share the flexibility in geometry or flow type that the penalty immersed boundary approach does, it is a much faster calculation and so allows for a more comprehensive exploration of possible hair arrangements over the cercus, while avoiding the overestimation of coupling inherent in our earlier effort. This model (Cummins 2009) predicts 50–66 % more damping in the motion of a short cricket hair in-line with long hairs than does the work of Bathellier et al. (2005, 2010). At this time, experiments are being carried out in Dr. John Miller's lab to determine the amount of coupling occurring between filiform hairs on cricket cerci.

4. Hair responses to biologically relevant stimuli

The ability to model biological arrangements of hairs and perturbations to those arrangements is important for understanding the specific functionality of sensor arrays in different species. It is likewise important to be able to predict the mechanical response of groups of hairs to biologically relevant signals. So far, most of the numerical modeling of filiform hairs and a fair amount of the experimental work has concentrated on oscillatory signals or band-limited white noise. Except in rare cases, this stimulus regime does not mimic natural conditions. In general, air velocity sensors are required to operate in noisy environments and to capture complex signals that may have transient or spatially rich components that indicate the presence of prey, predators, or mates. One of the next great challenges in filiform hair modeling is to increase the range of stimuli for which we can predict hair response.

The penalty immersed boundary approach of Heys et al. (2008) is already capable predicting the response to complex stimuli, although only computations using simple stimuli have been performed thus far. Outside of a full Navier-Stokes numerical solver, modeling stimuli that are spatially complex remains a challenge. Recently, Kant and Humphrey (2009) have succeeded to model hair response to very short, abrupt accelerations of an object. These accelerations induce airflow pulsations mimicking the beginning of an attack lunge of a spider (Casas et al. 2008), or of the vortices in the wake of a flying prey insect as it passes a spider (Barth and Höller 1999). The pulsations decrease in magnitude with both time from the event and distance from the source. Kant and Humphrey (2009) approximate the spatial decay as that of a spherical dipole, where the radius of the sphere is fit to empirical data from Dangles et al. (2006a) and Casas et al. (2008). For ease of calculation, they assume a Gaussian temporal decay of the pulse. With these choices, Kant and Humphrey then calculate the boundary layer arising from this spatially and temporally decaying pulsation source along the cercal (or spider leg) axis. With this method, they are able to predict the response of different hairs with varying position from the

lunging object to pulses with differing temporal structure. They find that shorter hairs respond more quickly and with larger magnitude to pulsations than do longer hairs in both spiders and crickets. This is perhaps not an unexpected result given the decreased inertia of the shorter hairs.

We can use the output of our model (Cummins 2009) to predict hair motion induced by temporally complex bulk flows by using a complementary approach to that of Kant and Humphrey (2009). Instead of characterizing the *stimulus* using Fourier methods, we characterize *hair mechanical response* using Fourier methods, specifically by calculating an approximate impulse response function of a hair. An impulse response function (IRF) is the response of a system to a Dirac-δ pulse. When convolved with an arbitrary signal, it predicts the response of the underlying system. In our case, it predicts the mechanical motion of a filiform hair subject to a mixture of driving airflows at different frequencies.

An IRF can be calculated by taking the inverse Fourier transform of the mechanical transfer function of a system, where the transfer function is the complex-valued function of frequency that encodes the gain and phase response of the system. Since a true IRF encodes the response of the system to all frequencies, it is never exactly recovered by any finite number of frequency measurements. Nonetheless, we may approximate the IRF of a given filiform hair by calculating the hair's response at each frequency up to the point where the response attenuates, and then taking the inverse Fourier transform of the truncated transfer function (see Cummins 2009 for details). The IRF will depend on the biomechanical parameters of the hair, on the number and arrangement of neighboring hairs, and on the appendage diameter at the hair base.

The approximate IRF serves as general response characteristic for a filiform hair, just as transfer functions traditionally have done. An IRF encodes the same information as a transfer function, but in the time domain rather than the frequency domain. The advantage of an IRF is that it may be convolved with any air velocity time trace to predict the hair's response to the signal. This convolution is equivalent to taking the Fourier transform of the signal, calculating the resultant boundary layer over the cercus for each frequency component in the truncated range, simulating the hair response to each frequency input, and summing all of the hair responses together. This latter approach is similar to the work of Kant and Humphrey (2009), except that the stimulus does not exhibit prescribed spatial or temporal decay.

Using IRFs, we characterized the responses of long (1 mm, ~8 µm base diameter) and short (0.25 mm, ~3 µm base diameter) filiform hairs with axial directions of motion. The hairs are either isolated or they have two neighbors with the same biomechanical properties as the reference hair. The two flanking neighbors are oriented either along the cercal axis or transversely across the 0.5 mm diameter cercus, as in arrangements 2) and 3) in Fig. 2 a. The axial separations are 0.3 mm (long hairs) or 0.12 mm (short hairs) and the transverse arc length separations are 0.1 mm (long hairs) or 0.07 mm (short hairs). The associated nondimensional spacings (distance/base diameter) are 37.5, 40.0, 12.5, and 23.3 respectively. These separations were the minimum found for long and short hairs near the proximal end of a representative *Acheta domesticus* cercus. The biomechanical parameters of the hairs and their basal diameters were calculated from regression relationships for *Gryllus bimaculatus* in Shimozawa et al. (1998). The transfer functions of the hairs were calculated out to 600 Hz and the IRF was calculated with a Nyquist frequency of 6000 Hz. The Nyquist frequency is half the sampling frequency in the time domain, so

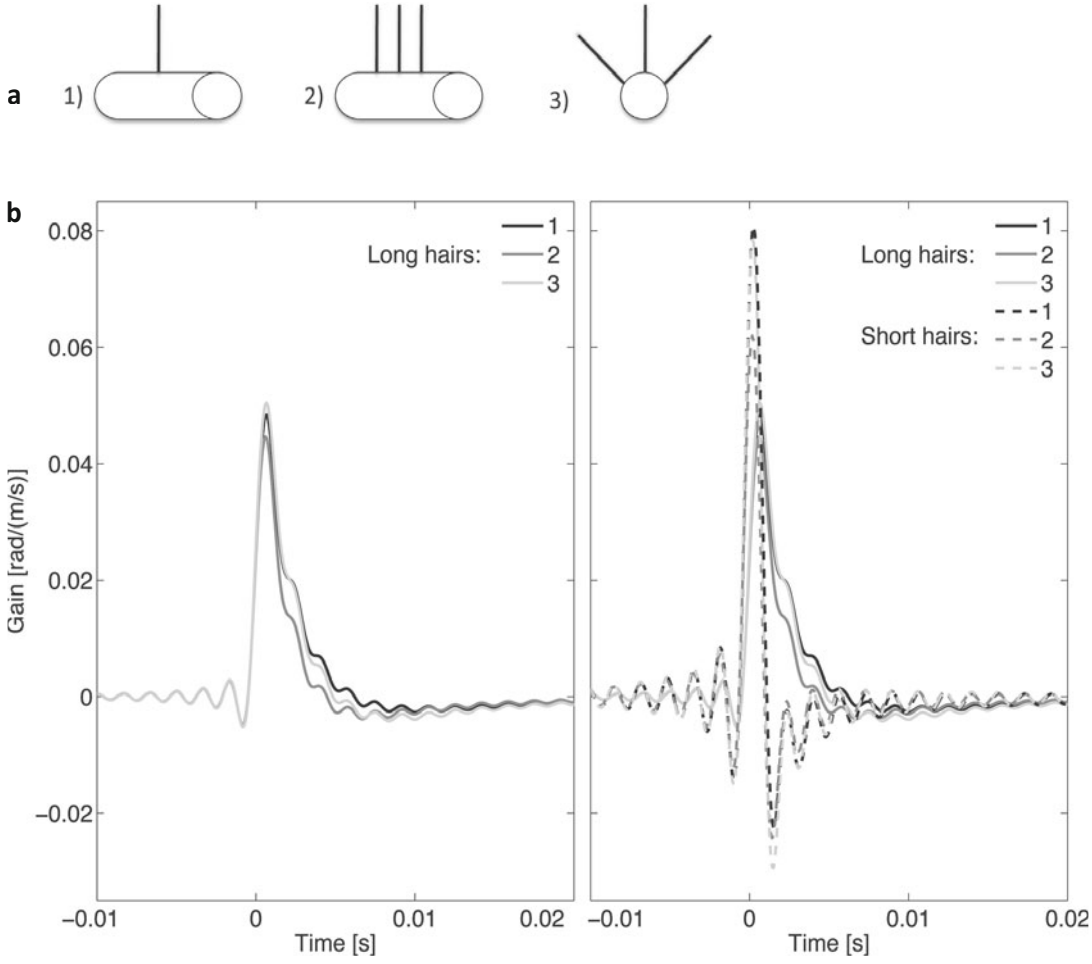


Fig. 2 **a** Schematic showing the hair arrangements for the impulse response functions (IRFs) plotted in panel b. Arrangement 1) is an isolated hair; 2) shows axial neighbors; and 3) shows transverse neighbors. Note that the schematic is not to scale. **b** Left panel: Plot of the approximate IRFs for a 1 mm hair that is isolated (black), has 1 mm neighbors with 0.3 mm axial separation (dark gray), or has neighbors with 0.1 mm arc length separation (light gray). Right panel: The solid lines are the same as in the left panel and are included in the plot for comparison purposes. The dashed lines are the IRFs for a 0.25 mm hair with and without 0.25 mm neighbors. Axial neighbors are at 0.12 mm separation and transverse neighbors are at 0.07 arc length separation

that the time step of the discrete approximate IRF is 1/1200 of a second.

The solid lines in Fig. 2b are the IRFs for the 1 mm hair and the dashed lines are those for the 0.25 mm hair, with black lines representing the isolated hairs and shades of gray representing the hairs with neighbors. The left panel in Fig. 2b displays the long hair response alone for clarity, and the right panel includes both long and short hairs for comparison between hair lengths. For the most part, the hairs with transverse neighbors do not differ much from their respective reference curves, although slight amounts of synergistic coupling can be seen at the largest peak in the long hair IRF and at the largest dip in the short hair IRF. The hairs with axial neighbors show damped IRFs throughout the majority of the trace. The comparison between hair lengths is quite striking, in

that short hairs respond to a δ pulse more swiftly and with greater magnitude than long hairs, even in the presence of neighboring short hairs. This result is in agreement with the work of Kant and Humphrey (2009), although the stimulus is simpler: It is an approximate δ pulse that induces a boundary layer both spatially and temporally uniform along the cercal axis. As mentioned earlier, this result agrees with an intuitive argument based on the differential inertia between long and short hairs. We are currently working on verifying this mechanical response on the cricket cercus and determining if the distribution of short hairs might confer selective advantage on the organism.

In order to use the IRF method to predict hair motion, it will be necessary to establish estimates of the decay in velocity magnitude for arbitrary stimuli, make appropriate assumptions about the formation of the boundary layer, and account for time lags in responses of hairs in a group. The interplay between coupling and complex stimuli remains a significant challenge in modeling total hair canopy response.

Conclusions and outlook

The goal of realistically modeling the mechanical response of a population of filiform hairs still faces significant challenges, but the advances in the last few years have provided a means for predicting responses to a large number of stimulus types. Immersed boundary method approaches are computationally intensive, but can model naturally occurring, complex stimuli, while unsteady Stokes methods are faster but less general. Exciting opportunities remain for combining these methodologies to leverage the generality of the immersed boundary method without incurring its full computational cost.

References

Barth FG (2002) Spider senses–technical perfection and biology. Zoology 105: 271–285

Barth FG, Höller A (1999) Dynamics of arthropod filiform hairs. V. The response of spider trichobothria to natural stimuli. Phil Trans R Soc Lond B 354: 183–192

Barth FG, Wastl U, Humphrey JAC, Devarakonda R (1993) Dynamics of arthropod filiform hairs. II. Mechanical properties of spider trichobothria (*Cupiennius salei* Keys.). Phil Trans R Soc Lond B 340: 445–461

Bathellier B, Barth FG, Albert JT, Humphrey JAC (2005) Viscosity-mediated motion coupling between pair of trichobothria on the leg of the spider *Cupiennius salei*. J Comp Physiol A 191: 733–746

Bathellier B, Barth FG, Albert JT, Humphrey JAC (2010) Erratum to Bathellier et al. (2005). J Comp Physiol A 196,1: 89

Casas J, Steinmann T, Dangles O. (2008) The aerodynamic signature of running spiders. PLoS ONE 3: e2116 (doi: 10 1371/journal.pone.0002116)

Cheer AYL, Koehl MAR (1987) Paddles and rakes: Fluid flow through bristled appendages of small organisms. J Theor Biol 129: 17–39

Cortez R (2001) The method of regularized Stokeslets. SIAM J Sci Comput 23: 1204–1225

Cortez R, Cummins B, Gregg KL, Varela D (2010) Computation of Brinkman flows using regularization methods. J Comp Phys A 229: 7609–7624

Cummins B (2009) Determining the biomechanical response of a filiform hair array: a low Reynolds number fluid-structure model. PhD thesis, Montana State University, Bozeman MT

Cummins B, Gedeon T (2007) A refined model of viscous coupling between filiform hairs in the cricket cercal system. Proc IMECE2007, #41075

Cummins B, Gedeon T, Klapper I, Cortez R (2007) Interaction between arthropod filiform hairs in a fluid environment. J Theor Biol 247: 266–280

Dangles O, Magal C, Pierre D, Olivier A, Casas J (2005) Variation in morphology and performance of predator-sensing system in wild cricket populations. J Exp Biol 208: 461–468

Dangles O, Ory N, Steinmann T, Christides JP, Casas J (2006a) Spider's attack versus cricket's escape: velocity modes determine success. Anim Behav 72: 603–610

Dangles O, Pierre D, Magal C, Vannier F, Casas

J (2006b) Ontogeny of air-motion sensing in cricket. J Exp Biol 209: 4363–4370

Fletcher NH (1978) Acoustical response of hair receptors in insects. J Comp Physiol A 127: 185–189

Gnatzy W, Tautz J (1980) Ultrastructure and mechanical properties of an insect mechanoreceptor: stimulus-transmitting structures and sensory apparatus of the cercal filiform hairs of *Gryllus*. Cell Tissue Res 213: 441–463

Heys J, Gedeon T, Knott BC, Kim Y (2008) Modeling arthropod hair motion using the penalty immersed boundary method. J Biomech Eng 41: 977–984

Humphrey JAC, Barth FG (2008) Medium flow-sensing hairs: biomechanics and models. In: Casas J, Simpson SJ (eds) Advances in insect physiology: Insect mechanics and control. 34: 1–80. Academic Press: London

Humphrey JAC, Devarakonda R, Iglesias I, Barth FG (1993) Dynamics of arthropod filiform hairs. I. Mathematical modelling of the hair and air motions. Phil Trans R Soc Lond B 340: 423–444

Kant R, Humphrey JAC (2009) Response of cricket and spider motion-sensing hairs to airflow pulsations. J R Soc Interface doi: 10 1098/rsif.2 008 0523

Kim Y, Peskin CS (2007) Penalty immersed boundary method for an elastic boundary with mass. Physics of Fluids 19: 053103–18

Kumagai T, Shimozawa TA, Baba Y (1998a) Mobilities of the cercal wind-receptor hairs of the cricket. J Comp Physiol A 183: 7–21

Kumagai T, Shimozawa TA, Baba Y (1998b) The shape of wind-receptor hairs of cricket and cockroach. J Comp Physiol A 183: 187–192

Landolfa MA, Jacobs GA (1995) Direction sensitivity of the filiform hair population of the cricket cercal system. J Comp Physiol A 177: 759–766

Magal C, Dangles O, Caparroy P, Casas J (2006) Hair canopy of cricket sensory system tuned to predator signals. J Theor Biol 241: 459–466

Osborne L (1996) Signal processing in a mechanosensory array: dynamics of cricket cercal hairs. PhD thesis, University of California, Berkeley

Shimozawa TA, Kanou M (1984a) Varieties of filiform hairs: range fractionation by sensory afferents and cercal interneurons of a cricket. J Comp Physiol A 155: 485–493

Shimozawa TA, Kanou M (1984b) The aerodynamics and sensory physiology of a range fractionation in the cercal filiform sensilla of the cricket *Gryllus bimaculatus*. J Comp Physiol A 155: 495–505

Shimozawa TA, Kumagai T, Baba Y (1998) Structural scaling and functional design of the cercal wind-receptor hairs of cricket. J Comp Physiol A 183: 171–186

Stokes GG (1851) On the effect of the internal friction of fluids on the motion of pendulums. Trans Cambridge Philos Soc 9: 8ff (Reprinted in Mathematical and physical papers Vol. III, pp. 1–141, Cambridge University Press, 1901)

Tautz J (1977) Reception of medium vibration by thoracal hairs of caterpillars of *Barathra brassicae* L. (Lepidoptera, Noctuidae). J Comp Physiol A 118: 13–31

Spider strain detection

17

Friedrich G. Barth

Contents

Abstract

This chapter is about biological mechano-sensors embedded in the cuticular exoskeleton of arthropods and serving purposes analogous to those of technical strain gauges. Spiders have several thousands of such sensors providing them with a highly resolved picture of the mechanical events in their exoskeleton. The spider "strain gauges", also called slit sensilla, form tiny slits (length ca. 8 to 200 µm, width 1 to 2 µm) in the exoskeleton, each supplied by two sensory cells. The slits are sites of locally increased compliance which respond to the slightest deformation of the exoskeleton caused by muscle activity, hemolymph pressure, or external sources of load such as substrate vibrations. Slit compression by as little as ca. 1.5 nm and due to strains down to some –10 to –20 µε are sufficient to set off action potentials in the sensory cells. A particularly intriguing feature of slit sensilla are lyriform organs representing diverse and sometimes seemingly bizarre close parallel arrays of up to ca. 30 slits.

The following text will highlight some of the properties of the sensors per se and relate their structural characteristics to natural sources of load and the strains produced by them in the exoskeleton. Considering the limitations of direct micromechanical measurements, the modelling of strains and organ deformation by Finite Element analysis has been particularly rewarding. Computational biomechanics has thus contributed substantially to revealing the principles underlying the measurement of strains in a highly refined biological sensory system. It also has paved the way for the future development of bio-inspired artificial strain sensors for technical applications.

Friedrich G. Barth
University of Vienna, Faculty of Life Sciences
Department of Neurobiology
Althanstr. 14, 1090 Vienna, Austria
e-mail: friedrich.g.barth@univie.ac.at

1. Introduction

The forces contained in the adequate stimuli of all mechanoreceptors generate some strain in the sensory structures. Indeed, stress, which is force per unit area ($\sigma = F/A$), and strain ($\varepsilon = l/l_o$), which is relative deformation due to stress, are inseparable from each other. The distinctive feature of the strain sensors treated here is their close relationship to the arthropod exoskeleton. The cuticular exoskeleton of arthropods like insects and spiders serves not only to mechanically support the animal and to provide the necessary leverage for locomotion, but also protects it from unwanted environmental influences, and may camouflage or alternatively advertise its bearer by adopting an appropriate shape and color. In addition, the exoskeleton represents the interface between the outside and inside of the organism and, therefore, carries a wealth of sensory structures on its surface that collect information about the outside world. Importantly, the exoskeleton serves its sensory function not only as a base for large numbers of exteroreceptors, but also as a stimulus conducting structure, which propagates stresses and strains to specialized sensors embedded in it and forming a skeletal sense unique to arthropods and alien to humans and other vertebrates. The slit sensilla of spiders form the most elaborate sensory system of this type and will, therefore, be the focus of our attention. Slit sensilla provide spiders with a detailed picture of the mechanical events going on in their exoskeleton which they use as guidance for various patterns of behavior.

Arachnid slit sensilla, like insect campaniform sensilla, are analogues of technical strain gauges. The wealth of strain gauges now commercially available reflects their general importance for many areas of engineering like the construction of buildings, bridges and industrial machines, cars, airplanes and rockets, to name just a few. Like strain gauges the slit sensilla measure the effects of force. In our present biological case these forces act on the exoskeleton and are prompted either by muscular contraction and hemolymph pressure (used to extend the leg), or to external sources like vibration of the substrate and gravity. Both the slit sensilla and campaniform sensilla are tiny holes in the exoskeleton and spots of enhanced compliance which slightly change their shape when the exoskeleton is minimally deformed by such forces. The sensory cells attached to the holes are excited by this deformation.

What follows will not only highlight properties of the sensors (sensilla) per se. An in-depth understanding of the principles at work and of the achievements made possible by their application also asks for the consideration of several components of the entire sensory system and the ways in which they influence, shape and indeed reflect each other. The focus here (albeit treated with different levels of detail) will be on the occurrence and topography of the sensilla, their structural characteristics, the natural sources of load, and the strains in the exoskeleton resulting from them. Having established this background the focus will then be on the mechanical tricks involved in the sometimes bizarre close parallel arrangements of slits forming so-called lyriform organs. The data gained from electrophysiological studies will highlight physiological properties of the sensory cells associated with the slit sensilla.

Finally, some conclusions will be drawn regarding the development of synthetic slit sensilla for technical applications. For earlier reviews on arthropod skeletal strain sensilla, see Barth (1981, 2002), Barth and Blickhan (1984), and Vincent et al. (2007).

2. Occurrence and topography

We shall concentrate here on a few basic statements only. The reader interested in more detail is referred to the literature (Barth and Libera 1970; Barth and Wadepuhl 1975; Barth and Stagl 1976). Throughout the chapter we will return to the question of sensilla topography several times, in particular in the context of sources of load and strain measurements at various sites of the exoskeleton.

Slit sense organs were first referred to by Bertkau (1878) but first shown to be mechanoreceptors only much later by Pringle (1955). They are a unique feature of arachnids and best studied in spiders. Whereas the number, arrangement and position of slit sense organs are very similar among spiders, their total number and degree of grouping (see below) differ considerably in scorpions (Barth and Wadepuhl 1975), whip

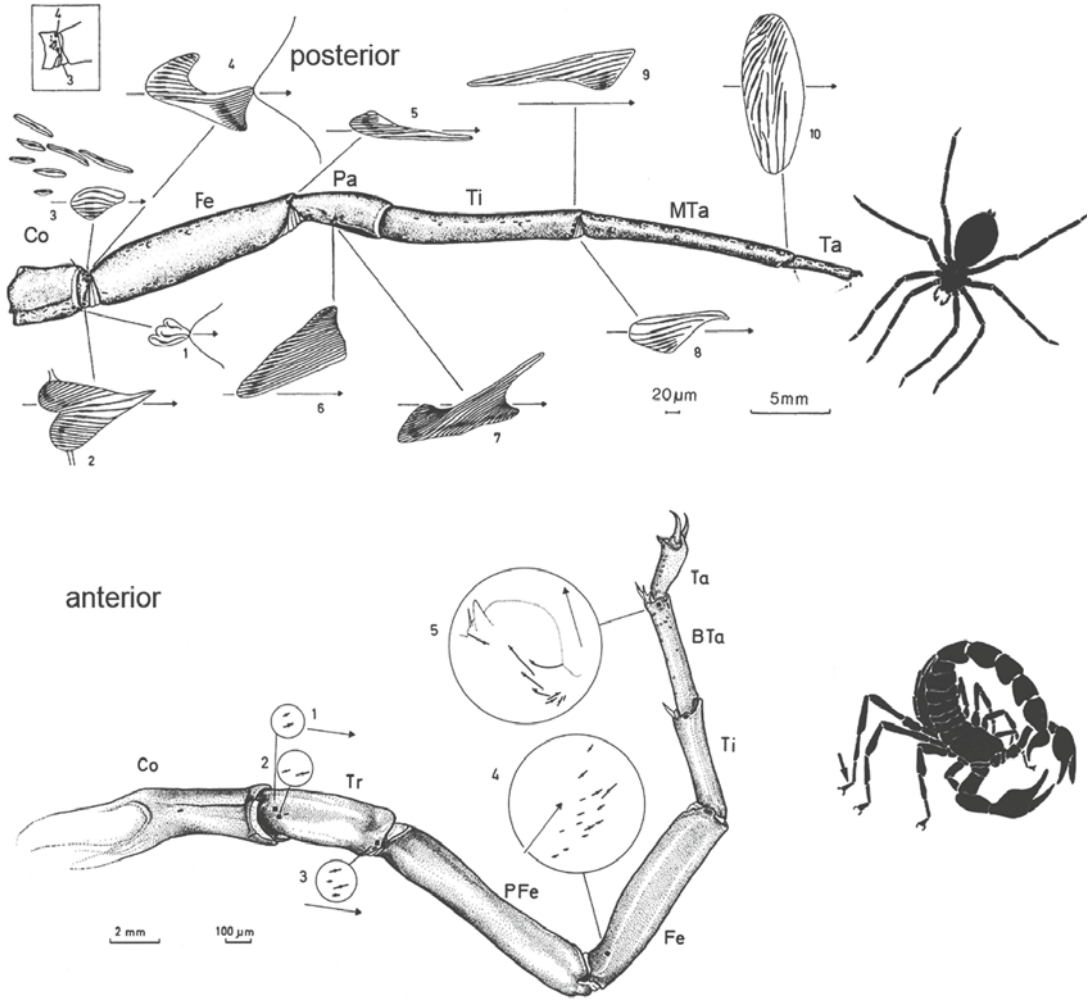


Fig. 1 *Occurrence and topography*. Above: spider leg (*Cupiennius salei*) and scorpion leg (*Androctonus australis*) showing lyriform organs and groups of slits. BTa, basitarsus, Co, coxa; Fe, femur; MTA, metatarsus; Pa, Patella; PFe, praefemur; Ta, tarsus; Ti, Tibia; Tr, trochanter. (from Barth and Libera 1970; Barth and Stagl 1976, with permission, Springer Heidelberg)

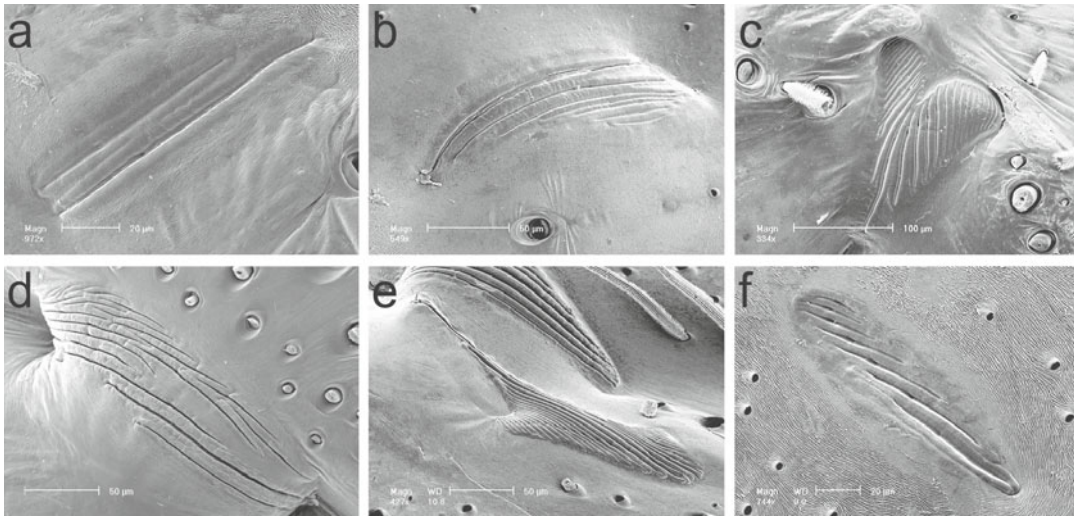


Fig. 2 *Lyriform organs.* Scanning electron micrographs of lyriform organs on the spider leg (*Cupiennius salei*). Organs shown in **b**, **c**, **d**, and **e** correspond to organs numbered 8, 2, 10, and 7 in Fig. 1 (with permission from Müllan 2005)

spiders, whip scorpions and daddy-long-legs (Barth and Stagl 1976). The whip spider and the daddy-long-leg we examined were the least well equipped. They had only 58 and 45 slits per walking leg as compared to 352 in the Central American wandering spider *Cupiennius salei,* which for many years has been our main experimental animal.

Cupiennius has about 3.300 sensory slits in its exoskeleton (Barth and Libera 1970). The majority of them (86 %) are embedded in the hard sclerotized cuticle of the walking legs and pedipalps. On the opisthosoma ("abdomen", excepting the petiolus) where they are surrounded by relatively soft mesocuticle, only 96 slits were counted. Arachnid slit sensilla occur in three types of arrangement: single slits (at least 100 µm away from the next slit), loose groups formed by several slits, and composite or lyriform organs representing conspicuous close parallel arrangements (slits separated by about 5 µm and never by more than 10 µm) of up to 30 slits. The distinction between these three types of slit arrangement was originally based on morphological criteria alone. However, it also has mechanical significance since the deformation of a slit by cuticular strains critically depends on the presence and distance of neighboring slits (as will be shown below). In *Cupiennius* one half of the slits compose a total of 144 lyriform organs, all of which are situated on the extremities, the majority (134) of them on the walking legs. From a biomechanics and biomimetics point of view, their strange patterns are of particular interest.

The majority of the lyriform organs are found close to joints (Fig. 1, 2), whereas single slits are arranged in rows on the front and back surfaces of the legs and pedipalps and sometimes close to muscle attachment sites. In both cases the slits are oriented roughly parallel to the limb's long axis. A few additional observations regarding the position of the slits on the body shall be mentioned here: *(i)* The slits are concentrated on the proximal end of the leg, where the stance phase of locomotion is generated and the musculature concentrated whose contraction causes strains in the exoskeleton. *(ii)* Most lyriform organs are located at

or very near to joints where forces are transmitted from one leg segment to the next and where a concentration of stresses is to be expected. *(iii)* Single slits are often found far away from joints but close to muscle attachment sites where bending forces are likely to occur. These rules are not exhaustive. Specific cases require careful examination of the particular area of the exoskeleton, as, for example, is illustrated most clearly in the spider vibration sensor (McConney et al. 2007; Fratzl and Barth 2009). We will also see that what matters for the stimulation of slit sensilla is not only the magnitude and direction of cuticular strains but also their time course.

3. Basic structure

Slit sensilla can be seen from outside through a dissecting microscope as fine elongated slits (about 8 to 200 µm long and 1 to 2 µm wide). In transmitted light their brightness indicates that they are holes in the exoskeleton covered by a thin (0.25 µm thick) and largely epicuticular membrane. The dendrite of one of the two sensory cells innervating each slit and picking up the stimulus (slit compression) ends in a specialized structure in this outer slit membrane which we refer to as the coupling cylinder (diameter ca. 0.5 µm, depth ca. 1 µm) (Fig. 3). The dendrite of the second sensory cell ends close to the membrane covering the inner side of the slit (Fig. 3 d).

These fine slits then are holes in the exoskeleton which are deformed by cuticular strains. The stiffness (Young's modulus) of the cuticle surrounding most of the slit sensilla is of the order of that found in bone (Blickhan and Barth 1979). The slits must, therefore, be very sensitive sensors and so they prove to be. Compression in the nanometer range suffices to set off nervous impulses in their sensory cells.

From an engineering point of view slits in a skeleton are problematic. They represent notches (openings) with a large length to width ratio (aspect ratio up to more than 100) where stress concentrations leading to cracks and fracture are to be expected (Peterson 1966). Nature takes care of these risks in at least three ways: (*i*) The resistance to pressure and tensile forces is increased close to the slits by a substantially increased thickness of the exocuticle around the opening, which contributes the most by far to the exoskeleton's mechanical strength (among its layers exo-, meso-, endocuticle) (Fig. 3 a). (*ii*) From a material's point of view the arthropod cuticle is a fiber reinforced laminate with the fibers arranged in layers parallel to the surface. Near the slits the fibers deviate from their arrangement parallel to the surface, following instead a course reminiscent of that of the stress lines near a notch subjected to tensile force or pressure. They curve inward like the lines of greatest principal stress conforming to the contours of the slit (Fig. 3 c). This means that the direction in which the load-bearing capacity of the cuticle is largest is the same as the direction of the greatest principal stress. (*iii*) The last trick dealing with the problem of notches under load, are the rounded (instead of sharp-edged) ends of the slits and fibers following their contours, moulded around them instead of being interrupted.

In nature there is no structure without function. In the given case morphological details of the slit become meaningful only in the light of stimulus transmission. Therefore, a short note is needed here to explain the primary mechanical processes. The following sequence of events characterizes the uptake of an initially diffuse stimulus from the general cuticle of the exoskeleton and

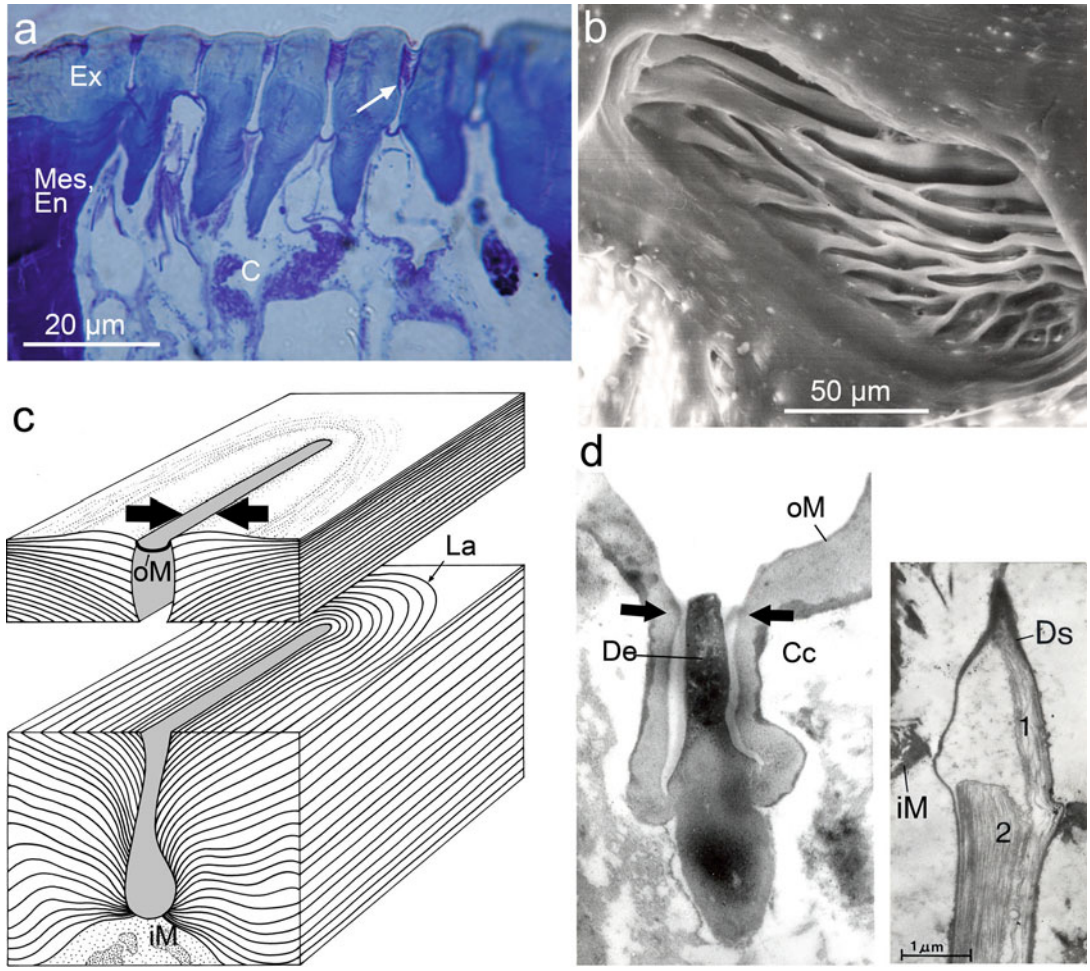


Fig. 3 *Basic structure*. **a** Cross section through the slits of a lyriform organ (*Cupiennius salei*). The slits with their outer and inner membranes can be clearly seen; Ex, exocuticle; Mes, mesocuticle; En, endocuticle; c, cellular components. **b** View of slits of the metatarsal vibration receptor organ (see also *d* in Fig. 2) as seen from inside an exuvia. **c** Arrangement of cuticular laminae around a slit according to fine structural analysis. oM, outer membrane; iM, inner membrane; arrows indicate adequate stimulation by compressional forces. **d** On the left side a cross section through the outer membrane (oM) is shown at the site of the coupling cylinder (Cc) which contains the dendritic ending (De) of one of the two sensory cells innervating each slit. The electron micrograph on the right shows the ending of the second sensory cell's dendrite (2) at the level of the slit's inner membrane (iM). The dendites of both sensory cells (1, 2) are contained in a cuticular dendritic sheath (Ds). (**a** and **c** with permission from Müllan 2005; **b** from Barth 2002 and **d** from Barth 1971, 2002, with permission, Springer Heidelberg)

its focus onto an area of only about 1 µm^2 of the dendrite: (*i*) According to both electrophysiological and biomechanical experiments (Barth 1972a, b) only compression of the slit triggers action potentials, not, however, its dilation. In a first approach it may be sufficient to say that the amount of slit deformation and its dependence on load direction increases with slit aspect ratio (length to width ratio). (*ii*) The slit's covering membrane is only 0.25 µm thick, supposedly considerably softer than the surrounding exo-

cuticle, and curved like a roof gutter which makes it mechanically reactive to compression of the slit; its curvature increases its bending moment in just the right direction (laterally imposed stress) (Fig. 3). (*iii*) The coupling cylinder is attached to the floor of the outer membrane where the bending moment due to slit compression is greatest. Its outer region is compressed mainly in a direction perpendicular to the slit long axis. (*iv*) Finally, there is a slight deformation of the dendrite tip by monoaxial compressive forces perpendicular to the long axis of the slit which changes the conductivity of the dendrite membrane and thus leads to a nervous response. The cuticular structures involved in stimulus transmission appear well adapted to being easily deformed. As in other mechanoreceptive sensors like spider hair sensilla (Barth 2004; Barth et al. 2004) the amount of deformation (in the given case that of the slit) is scaled down on its way to the dendrite which may be a protective measure. Whereas the basic mechanics of stimulus transmission seems to be quite clear for the dendrite ending in the outer slit membrane, that is not the case for the second dendrite ending at the inner membrane of the slit (Molina et al. 2009).

4. Sources of load and strains in the exoskeleton

As has been demonstrated in many cases (e.g. Wehner 1987; Barth 2004; Barth 2006; Fratzel and Barth 2009) sense organs to a large degree reflect the properties of their adequate stimulus, from the gross morphological to the micro- and molecular level. Therefore no sensor can be understood without sufficient knowledge of the stimulus patterns it is naturally exposed to and measures.

Key questions to be asked here are: Where do the loads/forces originate that lead to stress and strain in the cuticle and deforming the slits? What are the actual strains in the exoskeleton like? The first in-depth measurements of an arthropod that provided answers to these questions were those made by Reinhard Blickhan more than twenty years ago, with the spiders *Cupiennius salei* and *Aphonopelma* spec. (a bird spider which makes measurements easier due to its larger size). They refer to the legs where most of the slit sensilla are found and are still the only ones of their kind (Blickhan and Barth 1985; see also review in Barth 2002). Here again we have to concentrate on a few basic principles.

There are three main sources of load: (*i*) *Hemolymph pressure* which the spiders use to hydraulically extend their legs at joints lacking extensor muscles. (*ii*) *Contraction of the musculature* which is primarily found in the proximal section of the leg so that the forces of inertia are kept low as one would expect in a long and slender extremity moved rapidly and with large amplitudes. (*iii*) There are also *external sources* of load such as gravity and vibration of the substrate on which the spider is standing. Figure 4 summarizes the mechanical Bauplan of an arthropod leg which consists of thin-walled tube-like components connected by joints. Work concentrated on the tibia-metatarsus joint where four lyriform organs are found, among them the most extensively studied one (HS8, see Fig. 1). In addition the tibia-metatarsus joint does not have extensor muscles so that the effects of hemolymph pressure (used for extension) and muscle forces can be separated.

The force at the joint is dominated by the moments prompted by the hemolymph pressure and the muscle forces which act antagonistically. Although spiders have a so-called open circulatory system the blood flows in well defined spaces between the

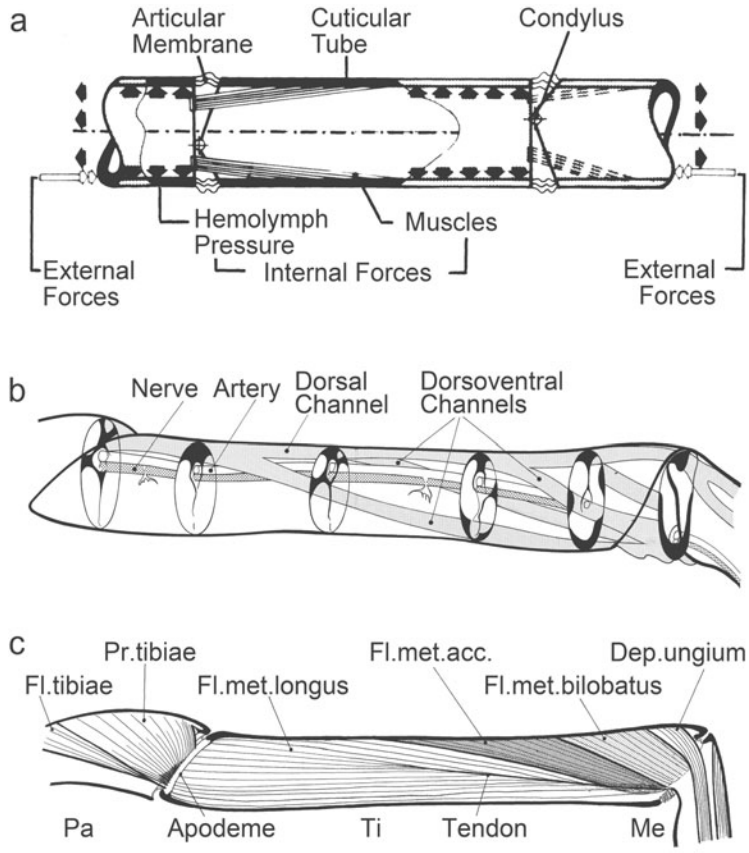


Fig. 4 *Sources of load*. **a** The legs of arthropods are composed of tubular segments connected by joints resulting in areas of contact much smaller than the segment's cross section. They are exposed to internal loads (filled arrows; hemolymph pressure; muscular force) and external loads (empty arrows; gravity; ground reaction force; substrate vibration). **b** Spider leg tibia showing the channels for the fast transport of hemolymph (blood) used to hydraulically move the joint. **c.** Arrangement of the flexor muscles in the spider leg tibia. Pa, patella; Ti, tibia; Me, metatarsus. (from Blickhan and Barth 1985, with permission, Springer Heidelberg)

muscles. Dorsoventral channels are of particular relevance here. They end at the greatest possible distance from the dorsal hinge joint. The result is a large moment $\vec{M}_h$ which is equal to the product of the pressure P and the effective area A on which it acts: $\vec{M}_h = P \times A\,(\vec{r}_s \times \vec{n}_A)$. Similarly, the moment $\vec{M}_m$ generated by the flexor muscles equals the product of the muscle force vector $\vec{F}_m$ and the position of the center of gravity of the muscle attachment area $\vec{r}_c$ with regard to the joint's axis of rotation: $\vec{M}_m = \vec{r}_c \times \vec{F}_m$. A third moment $\vec{M}_G$ contributing to the joint forces results from the ground reaction force $\vec{F}_G$ and its arms of force vector $\vec{r}_A$: $\vec{M}_G = \vec{r}_A \times \vec{F}_G$.

The mechanical sensitivity S_m = strain/force [ε/N] of exoskeletal areas with lyriform organs and at other locations of the tibia could be determined by applying miniaturized strain gauges at the respective sites and imposing controlled dorsoventral or lateral forces to the joint (using force transducer, displacement transducer, pressure transducer, and electrical stimulation of muscles). Muscular forces induce negative strain (compression) in the region of three of the four lyriform organs (HS8, HS9, VS4), but positive strain at organ VS5. In other words, the muscular activity effectively stimulates the three organs positioned laterally on the leg but not the ventral organ. Instead the ventral organ is stimulated (compressed) by hemolymph pressure. The mechanical sensitivity for *dorso-ventral forces* measured up to 20 με/mN which means that a 1 mN load changes the length of a piece of cuticle by 20×10^{-6}.

The mechanical sensitivity to *lateral forces* was only 10 με/mN near the lyriform organs and strongly depended on the joint angle. Finally, the mechanical sensitivity is smallest to *axial forces* with values below 0.8 με/mN and also highly dependent on the joint angle (Blickhan and Barth 1985).

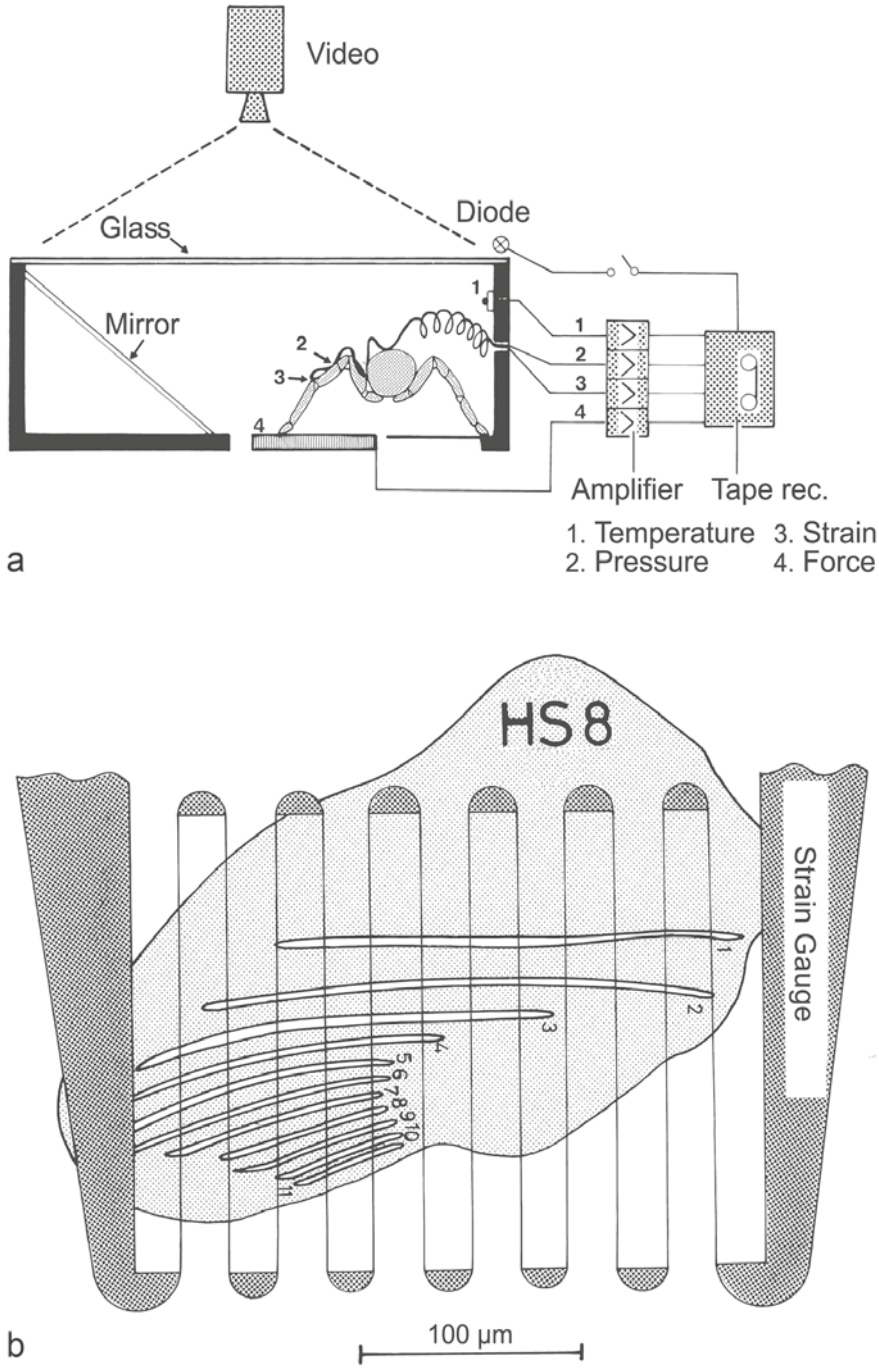


Fig. 5 *Direct measurements*. **a** Experimental set up allowing the simultaneous measurement of leg kinematics (video), cuticular strains (miniaturized strain gauges), ground reaction force (force platform), and hemolymph pressure (pressure transducer). In the walking track there is a mirror allowing to view the spider from the side and from above. **b** In vivo strain measurement at the site of a lyriform organ using a miniaturized strain gauge glued onto the cuticle of a lyriform organ (HS8) on the spider leg. (from Blickhan and Barth 1985, Barth 1985, with permission, Springer Heidelberg)

Although knowing the mechanical sensitivity to different loads allows for the prediction of strains occurring in a freely moving animal and thus the stimulation of lyriform organs, it cannot replace in vivo measurements. These require a lot of technology (Fig. 5) because many parameters have to be monitored simultaneously: stepping pattern, joint angle and leg kinematics (video), cuticular strains at the site of the lyriform organs (miniaturized strain gauges), hemolymph pressure (miniature pressure transducer), and the ground reaction force exerted on the ground by the legs (force platform).

Under conditions of slow walk (1 to 10 cm/s; 0.3 to 1.5 steps/s) and given the dependence of the mechanical sensitivity S_m on the joint angle (tibia-metatarsus joint) the strains in the spider leg tibia result mainly from muscular force (up to 200 mN) and hemolymph pressure (up to 5.3 kPa). Because of the spatial limitations on a spider leg, bird spiders were used for these experiments. Their legs are slightly larger in diameter than those of *Cupiennius.*

The magnitude of the cuticular strains found at the sites of the different lyriform organs close to the Ti-MeTa joint are very similar (Blickhan and Barth 1985) (Fig. 6a), varying only between –13 and –20 με. However, they change with the rhythm of the spider's movement and differ in time course at the different positions in relation to the walking pattern. Whereas organs HS8, HS9 and VS4 laterally on the tibia back and front side signal muscle force during the flexion of the joint (stance or load bearing phase, negative strain), the ventral organ VS5 is stimulated by hemolymph pressure during joint extension (swing phase) when the other organs are not compressed but dilated (positive strain). Lyriform organs are not located at sites of particularly high strain. Instead the significance of the differences in the time course taken by the cuticular strains at the different sites is underlined by the finding that on the dorsal surface of the leg, where there are no lyriform organs, strains as high as 38 με parallel to the leg long axis were found. Whereas the time courses of the strains are very similar in all four walking legs the strain amplitudes at the site of a particular organ increase from leg one to leg four (e. g. up to –80 με at the site of organ HS8 as compared to ca. –20 με on the second leg). This difference is mainly due to the flatter contact of the fourth leg with the substrate than that of the other legs, leading to an increase of the vertical component of the standing force (40 mN) (Blickhan and Barth 1985). Both dorsally, in the middle of the tibia, and at the insertion of the flexor muscle, strains measure up to –70 to –80 με during the stance phase.

This short summary underlines the importance of knowledge of the stimuli to be expected under natural conditions. It should be remembered that the exoskeleton is exposed to many forces which it can tolerate not only because of the cuticular material it is made of and its structural organization but also because mechanical loads can be detected and if necessary avoided by the many slit sensors embedded in it. The measurements described above were, therefore, extended by taking into account situations where one would expect particularly high strain values, namely during rapid starts and stops and jumps (Brüssel 1987; review in Barth 2002).

As the walking speed increases (10 to 30 cm/s) the load does so as well, especially in the joint region, where at 30 cm/s strains as high as –360 με were measured (stance phase). As long as the substrate is horizontal it does not make a difference whether the spider is walking upright or hanging from a surface. However, walking on a vertical plane is common for *Cupiennius* and other spiders. When walking upwards load increases 2- to 3-fold and strain values in the joint region are conspicuously irregular with similar strain

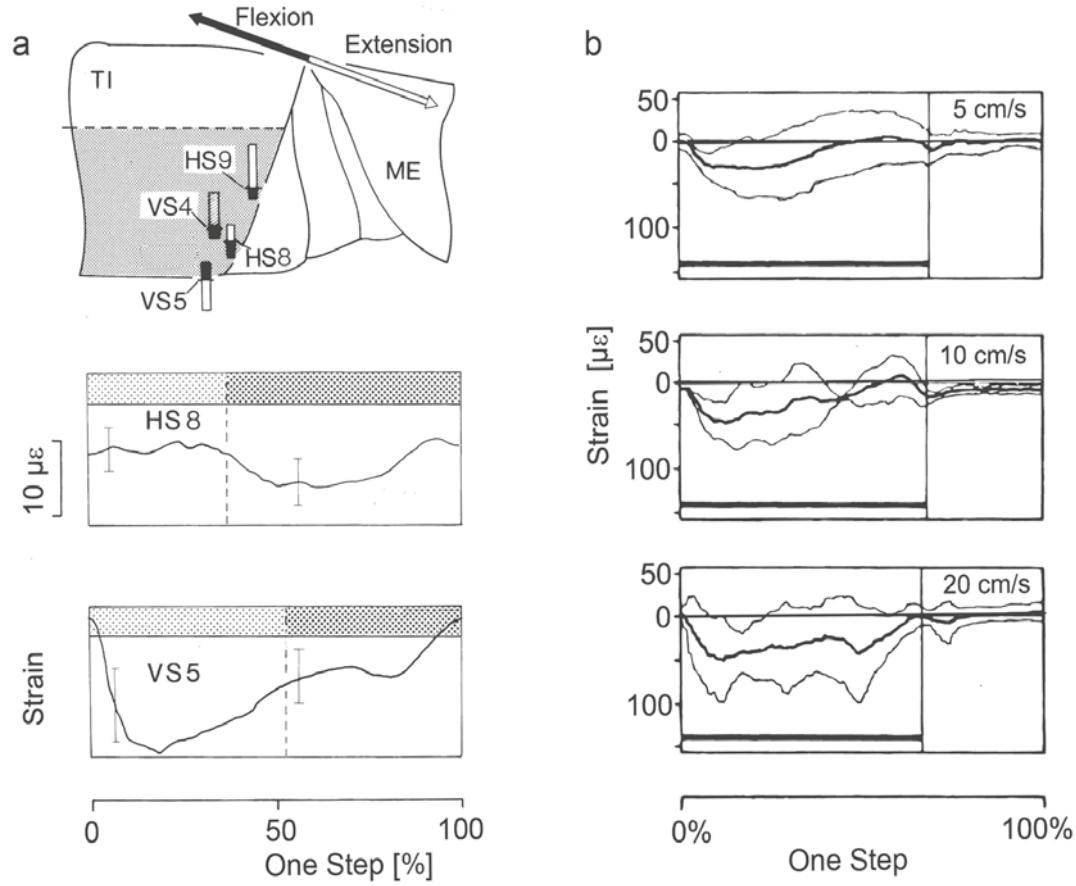


Fig. 6 *Strains in the exoskeleton*. **a** During flexion of the metatarsus (ME) by muscular contraction there are negative cuticular strains (see black bars below horizontal line indicating location of organ) leading to slit compression laterally on the leg tibia at the sites of organs VS4, HS8 and HS9 whereas ventrally (see organ VS5) there are positive strains (black bar above horizontal line), implying dilation of the slits. Note, that hydraulic joint extension by hemolymph pressure leads to effective stimulation of the ventrally located organ (VS5) only. The two graphs below the scheme show the mean strains (± SD) measured at organs HS8 and VS5, respectively, during one step of a freely walking spider. Stance and swing phase are marked by the dark and light shading of the bar above graph. **b** Strains (mean ± SD) measured laterally on the tibia (close to lyriform organs HS8 and HS9) in spiders (*Cupiennius salei*) walking upwards on a vertical plane at different speeds. (**a** from Blickhan and Barth 1985; **b** from Barth 2002 after Brüssel 1987, with permission, Springer Heidelberg)

magnitudes for all walking speeds (Fig. 6b). Downward walks at a speed of 20 cm/s loaded the joint region much more than walks in the horizontal plane (Brüssel 1987; Barth 2002). During starts and stops and jumping movements the high acceleration imposes loads similar to those measured during the most rapid walks.

Under the conditions of slow and uniform movement at least the tibia of the spider leg is an object of equally distributed strength, but this is not the case during rapid movements when loading concentrates in the distal joint region. This implies that the lyriform organs are in a region with strongly varying strains and high peak loadings. The lyriform organs are indeed well adapted to monitor such variable load conditions and to

avoid dangerous peaks by eliciting synergic reflexes (Seyfarth 1978, 1985), due to their position, orientation and physiological properties. As will be seen in the next section the close parallel arrangement of the slits in lyriform organs leads to a remarkable mechanical range fractionation, that is to say great differences in mechanical sensitivity among the slits and a large working range of the entire organ.

Before proceeding in more detail, an issue of potential confusion shall be addressed. Do the slit sensors respond to strain, to displacement, or to force? Are they force or strain receptors? There is no short and clear cut answer to this because it depends on the level at which it is sought: (*i*) At the sensor level a slit measures slit face displacement. Accordingly the slit is a displacement (or deformation) receiver. Whereas strain is the *relative* change in length, the slits measure absolute changes of length which are the sum of many unit strains or displacements. The deformability of the stimulus transmitting structures as described in section 3 emphasizes the importance of displacement as does the lack of closeness of the dendrite tip to the areas of particularly high stress along the circumference of the slit. (*ii*) When considering the slit as part of a larger piece of exoskeleton it may be called strain receiver, measuring local deformation of the skeleton, analogous to what technical strain gauges do. (*iii*) Following up this reasoning even further, a sensory slit is to be regarded as part of an entire leg (or another part of the body) and as a device which measures forces that cause the deformation of the exoskeleton. In this context, then, a slit is a force sensor.

A more unequivocal answer to our question emerges if we ask which parameter the spider will be most interested in. It will not be interested to know in isolation how the slit faces are displaced with respect to one another. Quite differently, it is important for the spider to know about the deformation of larger parts of the exoskeleton signalled by the respective strains in order to avoid fracture or to deduce from them information about the forces acting on the skeleton.

5. Modelling strains and organ deformation

The most peculiar form of slit sensilla are the lyriform organs. They come in arrangements that at first sight appear bizarre if one hypothesizes that they reflect mechanical properties specifically adapted to the strains occurring at a particular position in the exoskeleton (an assumption justified by a biologist's evolutionary thinking). Direct measurements of the micromechanics of the original sensors are limited regarding spatial resolution. Modelling as an alternative approach has helped greatly to achieve a better understanding of the principles underlying lyriform slit arrays.

Lyriform organs are arrays of individual slits. The first question then is what the mechanical properties of a single slit are. In early model studies the deformation of slits cut into araldite disks was measured under load applied in the plane of the disk (Barth 1972a, b). It could readily be demonstrated that the magnitude of slit deformation and its directionality strongly increased with slit aspect (length-to-width) ratio and absolute length. Likewise the slits were much more easily deformed by forces perpendicular to the slit long axis than by forces parallel (or intermediate) to it. In a single slit, deformation increases strongly towards the middle along the long axis. Thus, in order to achieve high mechanical sensitivity a single slit should be long and oriented perpendicular to the lines of the principal stresses. In ad-

dition the dendrite tip in the outer slit membrane should be located in the middle along the long axis of the slit.

In a more recent attempt (Hößl et al. 2006) to quantify the mechanical properties of slit sensilla, Kachanov's analytical approach (Kachanov 1994; Gorbatikh and Kachanov 2000), which was developed within fracture mechanics to calculate stress intensity factors at the tips of interacting cracks, was applied to study the deformation of the slit faces by far-field strains. Apart from understanding the mechanical behavior of lyriform slit sensilla the goal was also to provide a solid basis for the quantitative prediction of the deformation pattern of bio-inspired synthetic strain sensors for technical applications. Kachanov's method turned out to be appropriate for studying single slits and loose groups of slits (crack spacing more than about one half of crack length), not however for the study of lyriform organs where the distance S between slits typically is much smaller than half of slit length l_o, $S < 0.5\,l_o$ (Fig. 7a) (Hößl et al. 2006). The amount of lateral spacing between the slits strongly influences their mechanical interaction. According to the results of a Finite Element (FE) analysis, the errors for the crack opening distances along the slit are too large to be acceptable for distances between slits of $S < 0.25\,l_o$ (l_o being the length of the longest slit in the array) (Fig. 7a). Hößl et al. (2006) discuss extended versions of Kachanov's theory to overcome this difficulty but conclude that the most economic route for the analysis of slits as closely arranged as in lyriform organs (S typically $< 0.05\,l_o$) is FE analysis. FE analysis allows the study of a wide variety of slit shapes, positions and orientations in a lyriform array, of curved slits as they actually occur, and of slits in spatially curved regions of the exoskeleton.

As a continuation of initial work which applied simple direct measurements of the deformation of slits cut into polymer disks (Barth et al. 1984), planar FE models were developed using ABAQUS (ABAQUS Inc., Pawtucket RI) to calculate the relative slit face displacement of single slits and arrays of five interacting slits arranged in patterns typical of actual lyriform organs under uni-axial compressive far-field strains (Hößl et al. 2007). The models took the form of planar square discs measuring $400\,l_o$ by $400\,l_o$ with the slits at their center (l_o, length of the longest slit in the array). With these dimensions, boundary effects of slit face deformation were avoided. Meshing with 6-noded triangular plane stress elements was strongly refined in the vicinity of the slits (Fig. 7b) (Hößl et al. 2007). The models did not have an absolute length scale so that the results can be scaled to any slit size. In a first step slit face displacements D_c were calculated for the slit centers (c) and normalized with respect to the deformation D_{sc} of a single slit (sc) of aspect ratio $l_o/b = 100$ subjected to uni-axial far-field strain of $-25\,\mu\varepsilon$ (-2.5×10^{-5}) normal to the slit faces. For details of the methods and their biological justification see Hößl et al. (2007).

Based on the diversity of arachnid slit sensilla the effects of the following geometrical parameters were studied quantitatively: aspect ratio (l/b) of the slit, slit shape (capped rectangular or elliptical), geometry of slit's center line (straight, C- or S-shaped), load direction (Φ), lateral distance (S) between slits, longitudinal shift (λ) between slits, and slit length difference (Δl) of neighboring slits.

5.1 Single slit (uni-axial normal load)

Slit face displacements are most sensitive to changes in slit length l and load direction Φ, increasing with increasing l and an angle Φ of up to 90°. Interestingly, aspect ratio is of little influence only for values between 20 and 100 (or higher) typical of lyriform organs. It deserves particular attention in

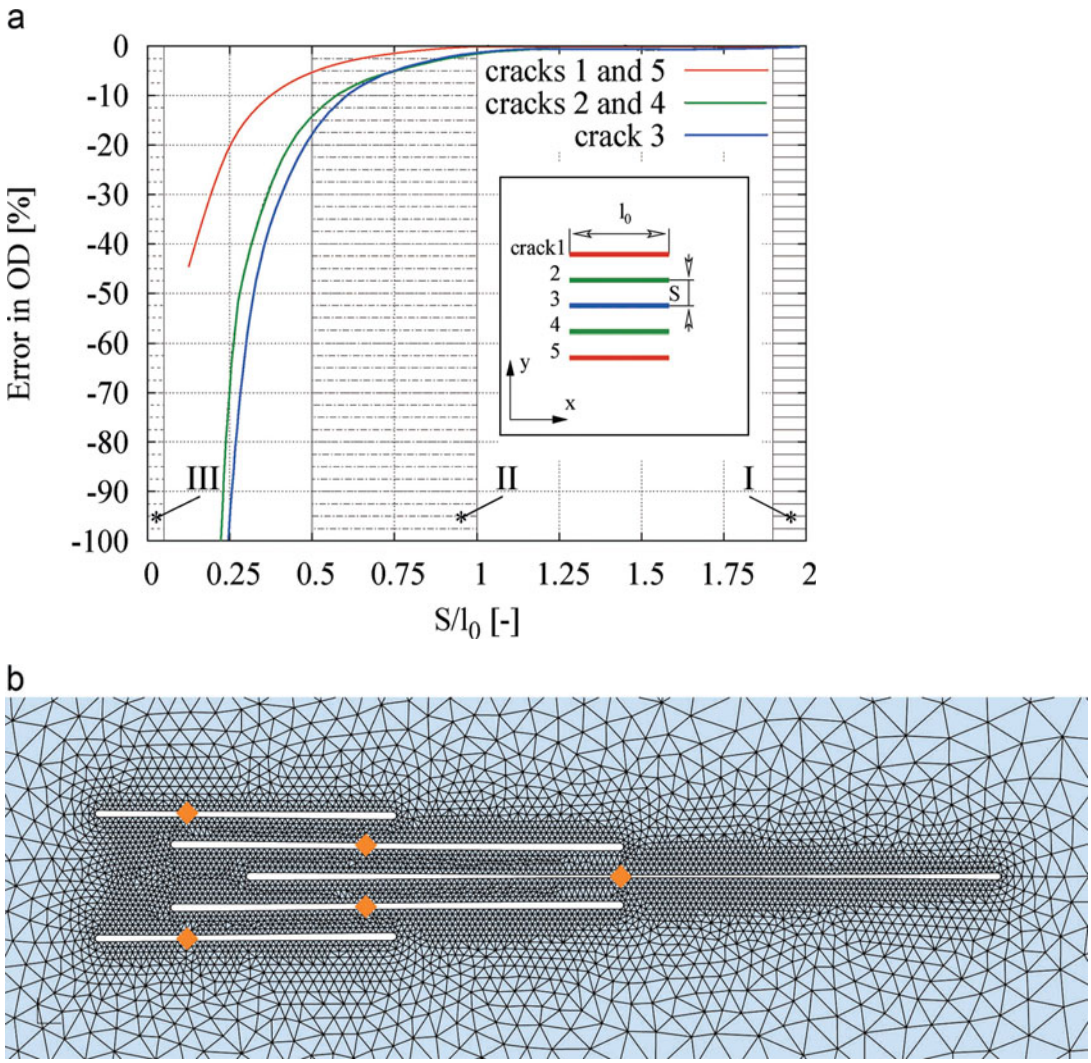


Fig. 7 *Modelling*. **a** A comparison of the opening deformation (*OD*) of 5 slits of length l_o lying in parallel as calculated by a fracture mechanical approach (b_k) (Kachanov's method; see text) and finite element (FE) analysis (b_{fe}), respectively. With decreasing distance *S* (and thus S/l_o) between the slits the error $(b_k - b_{fe})/b_{fe}$ in Kachanov's approximation increases. The ranges of spacing denoted by I, II, and III occur in slit sensilla defined as single slits, loose groups, and lyriform organs, respectively. **b** Finite element model of a lyriform organ ("heart shaped", compare Fig. 2c) with $S/l_o = 0.04$ deformed by load at 90° with regard to the slit long axes. For better visibility the displacements are scaled up by a factor of 10^4. Note the difference in magnitude of deformation between slits. Orange diamonds indicate sites of maximum slit face deformation. (**a** after Hößl et al. 2006 with permission Elsevier Ltd.; **b** after Hößl et al. 2007, with permission Springer Heidelberg)

regard to a comparison with the round to elliptical insect campaniform sensilla, where it rarely exceeds a value of 3 (Barth 1981). In Fig. 8a (Hößl et al. 2007) the deformation of a capped rectangular slit and of an ellipse, both of length l_o, are compared for aspect ratios between 1 and 100. For an aspect ratio of 1 the value for D_c/D_{sc} increases by 50 % (from 1 to 1.5). This is mainly due to the area of circular holes of diameter l_o which is

larger by about 80 % than that of a slit with an aspect ratio of 100. For all non-trivial aspect ratios capped rectangles (like in arachnid slits) show larger face displacements than ellipses (like many insect campaniform sensilla) of the same aspect ratio because of their larger area. The strong dependence of D_c/D_{sc} on the orientation angle Φ of the applied load results in maximal displacement at $\Phi = 90°$ (Hößl et al. 2007).

5.2 "Lyriform" arrays of five parallel slits

5.2.1 Uni-axial normal load ($\Phi = 90°$)

In the close parallel arrangements of slits shown in Fig. 7b the effects of lateral spacing, longitudinal shift, and length gradation were examined under uni-axial normal load. Whereas the lateral spacing S was changed when studying its effect on slit interaction in the non-staggered array, S/l_o was kept constant at 0.04 in the oblique bar and triangular arrays, which corresponds to the spacing found in real lyriform organs. FE modelling provided the predictions partly shown in Fig. 8c and d. Clearly, considerable amplification and shielding of the individual slits' face displacements can be achieved by varying S, λ, and Δl. Our studies of the effect of varying these geometrical parameters not only revealed functional principles of existing lyriform organs. They also allow for a reasonable hypothesis concerning their respective relevance during their evolution and how they contributed to adjust the organs to particular measurement demands/tasks like working range and resolution regarding stimulus magnitude and direction (see also further below).

(*i*) *Lateral spacing* in a non-staggered array of five slits has practically no effect on slit deformation as long as $S/l_o \geq 3$ where there is practically (< 3 %) no difference in D_c of neighbouring slits anymore. However, for slits as closely spaced as in natural lyriform organs ($S/l_o \leq 0.1$) the inner slits exhibit much smaller relative displacement D_c/D_{sc} than the shielding outer slits.

(*ii*) The effect of *longitudinal shift* λ between neighboring slits was studied in an oblique bar array with constant lateral spacing of S/l_o=0.04 (Fig. 8d). The relative slit face displacements are considerably larger than that of a single slit if $0.25 \leq \lambda/l_o \leq 2.5$; the largest amplification factor is predicted to be ca. 4.25 (425 %) for the center slit at $\lambda/l_o \approx 0.85$. If, on the other hand, λ approaches l_o, D_c decreases rapidly. Thus all slits of oblique bar arrangements show amplification and a similar level of face displacement. This characteristic potentially allows an enhancement of the signal-to-noise ratio by central convergence of the information provided by the individual slits. A longitudinal shift between neighboring slits is typical of the natural case, as is the gradation of slit length which implies a shortening of the cuticular ligaments between the slits and a decreasing effect of ligament bending (see also section 5.3.).

(*iii*) In triangular arrays with *slits of different length*, the shorter slits (numbers 2 to 5 in the array) are strongly shielded at length gradations $\Delta l/l_o \geq 0.1$ (Δl, difference in length between neighbouring slits; l_o length of longest slit in array) while the longest peripheral slit (slit 1) behaves like a single slit for $\Delta l/l_o \geq 0.09$ (Hößl et al. 2007). The result is a widely differing response of the slits in the array, suggesting that triangular configurations of slits serve to maximize the range of signal magnitude the organ can deal with. Among arachnid lyriform organs (Fig. 1 and 2b) length gradations vary between ≈ 0.08 and 0.3.

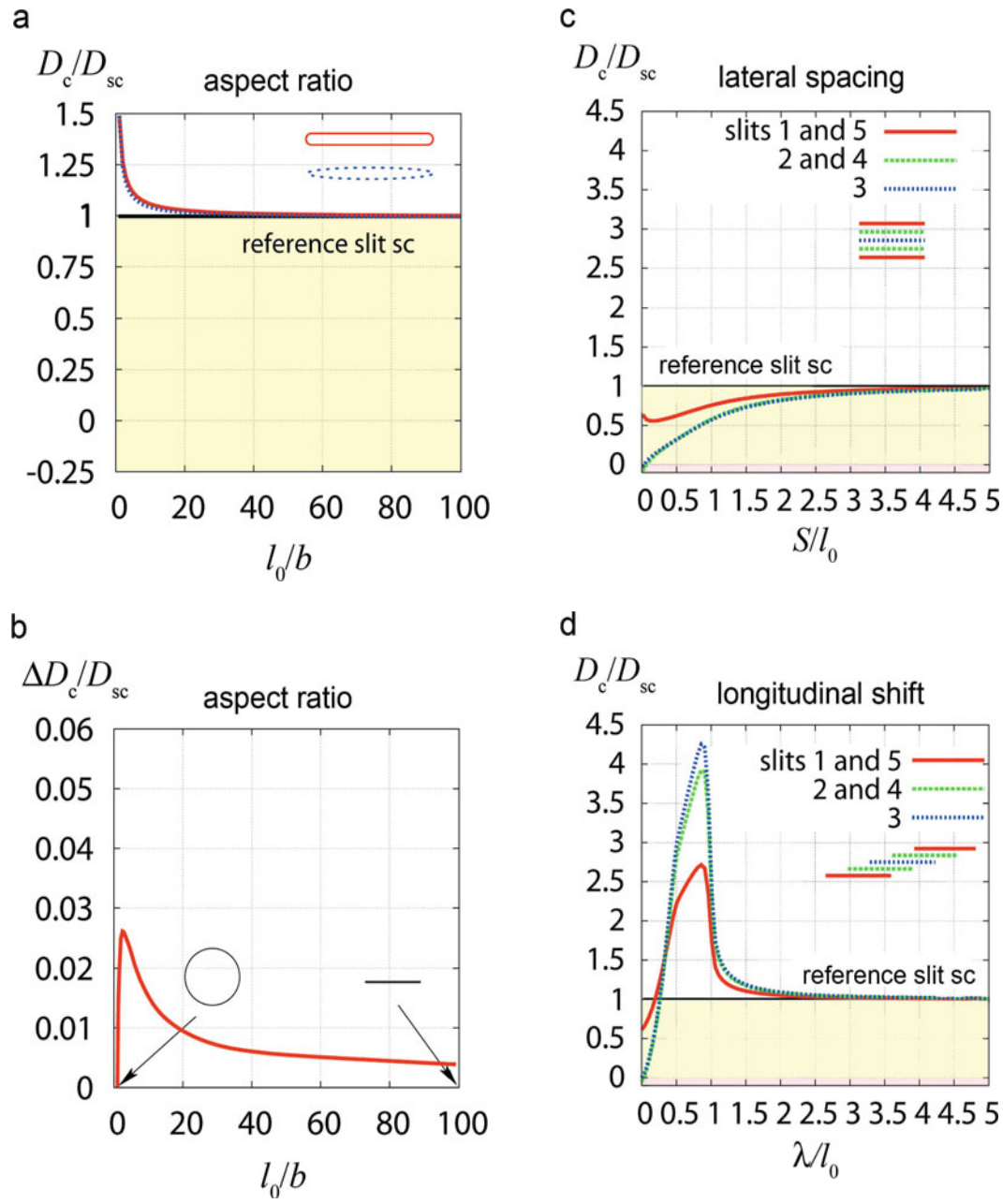


Fig. 8 Examples showing the influence of various geometric parameters of a single slit (**a**, **b**) and of an array of five parallel slits (**c**, **d**) on the normalized slit-face displacement at mid-length under load normal to the slits' long axes (*Dc/Dsc*, slit displacement relative to that of a single isolated slit). l_o, slit length; *b*, width of slit. **a** and **b** Effect of change of aspect ratio l_o/b on the deformation of a "capped rectangular" and an elliptical slit, both of length l_o. At an aspect ratio of 1 *Dc*/*Dsc* increases by 50 %, whereas at ratios > 20 the effect is considerably smaller. **c** and **d** Effect of lateral spacing (S/l_o) and longitudinal shift (λ/l_o) between five parallel slits on slit face deformation. Note considerable shielding and amplification. (from Hößl et al. 2007, with permission, Springer Heidelberg)

5.2.2 Dependence on load direction

As expected from the strong effect of load direction on the deformation of a single slit there are similar effects in lyriform arrays. Again maximum slit face displacement occurs at Φ close to 90°, the smallest or none at Φ ≈ 0°. This applies to the oblique bar and the triangular arrangement as well (Fig. 9). The angular range of mechanical sensitivity can be drastically enlarged by changing the slit configuration to heart-shaped or fan-shaped. The actual occurrence of such configurations suggests that they are adapted to different measurement tasks and/or differences in the directional pattern of the strains of interest.

5.3 Arrays closely matching natural lyriform organs

Carrying the FE analysis one step further, we approached the natural geometries of lyriform organs more closely than in case of the planar models of five slits arranged in different typical patterns (Hößl et al. 2009). Models were based on the details of actual organs (such as number of slits, lateral and longitudinal shift, curvature and the angle between the slits) and thereby extended the quantitative basis for the design of bio-inspired technical strain sensors. The overall result of this effort is that even minor variations in structural parameters may significantly change a slit's directional response (Fig. 10). Due to the quantitative assessment of the load conditions to which the models were subjected the slit deformation predicted by the FE-models could be meaningfully compared with the values in an intact natural organ found by taking direct measurements using white light interferometry. The agreement between the predictions of the model studies and the direct measurements was surprisingly good. This in turn made it possible to produce well-founded quantitative statements about the actual deformation of natural slits under biological strain conditions (see further below).

The oblique bar array will serve as an example to illustrate this point (see Hößl et al. 2009). As long as the slits are equally long their deformation is practically the same in this arrangement (Barth et al.1984, Hößl et al. 2007). This is not so in more realistic models (Fig. 10), where neither slit length nor longitudinal shift and angle between slits are constant. Whereas the overall directionality of the organ is the same as is the case for the simpler five-slit model (Fig. 9a), the graduation in mechanical sensitivity among the seven slits differs markedly from it. Note that whereas the upper model shown in Fig. 10 is already close to the natural slit morphologies (straight slits) the lower model organ copies the original slits even more precisely (slightly curved slits varying in width; based on scanning electron micrographs). The seemingly minor "improvement" of the model geometry significantly influences slit face deformation, the most obvious effect being a much more even distribution of the mechanical sensitivities of the shorter slits 4 to 6. Slit 7 is the smallest and one of the peripheral slits, shielded on one side by the larger slit 6. As one would expect its deformation is very small. In the original organ it will respond to very high strains/loads only.

5.4 A comparison of computational FE models with direct measurements

One lyriform organ has been studied more widely than any other with regard to the functional aspects of interest here. This is the "triangular" lyriform organ HS8 on the posterior aspect of the tibia close to the tibia-metatarsus joint, which in *Cupiennius salei* consists of 7 to 8 slits (Fig. 1 and 2b). It has previously been studied electrophysi-

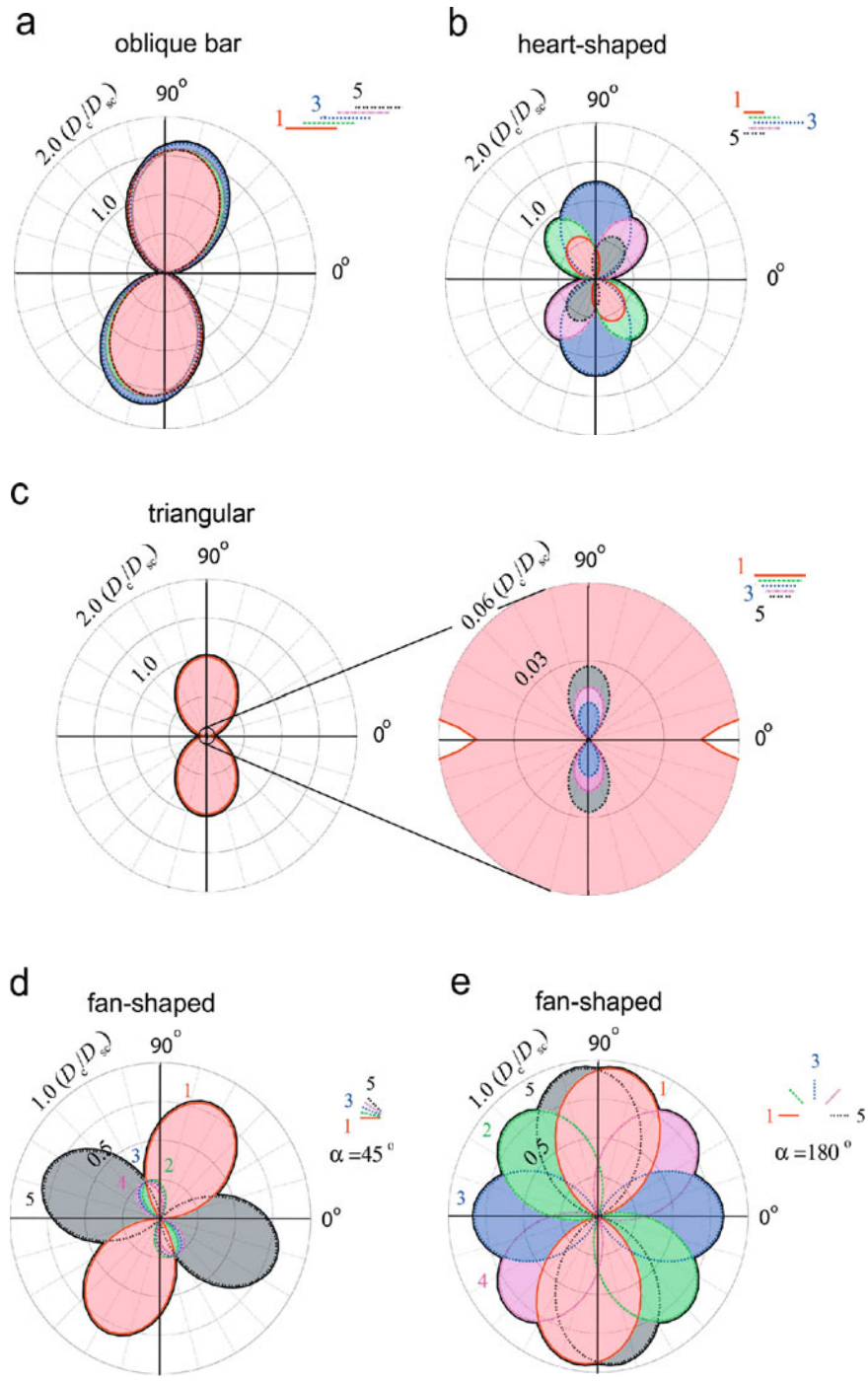


Fig. 9 *Directional sensitivities* to uni-axial compressive far-field stress of different "lyriform" arrangements of slits. **a**, **b**, and **c** In these model arrangements referred to as oblique bar, heart-shaped and triangular S/l_o was 0.04, simulating the natural case. Note differences in absolute sensitivity, range fractionation between slits, and range of directional sensitivity covered by the ensembles. **d** and **e** Effects of varying the angle between the slits' long axes. The directional range of mechanical sensitivity is considerably enlarged by increasing the angle between slits. (from Hößl et al. 2007, with permission, Springer Heidelberg)

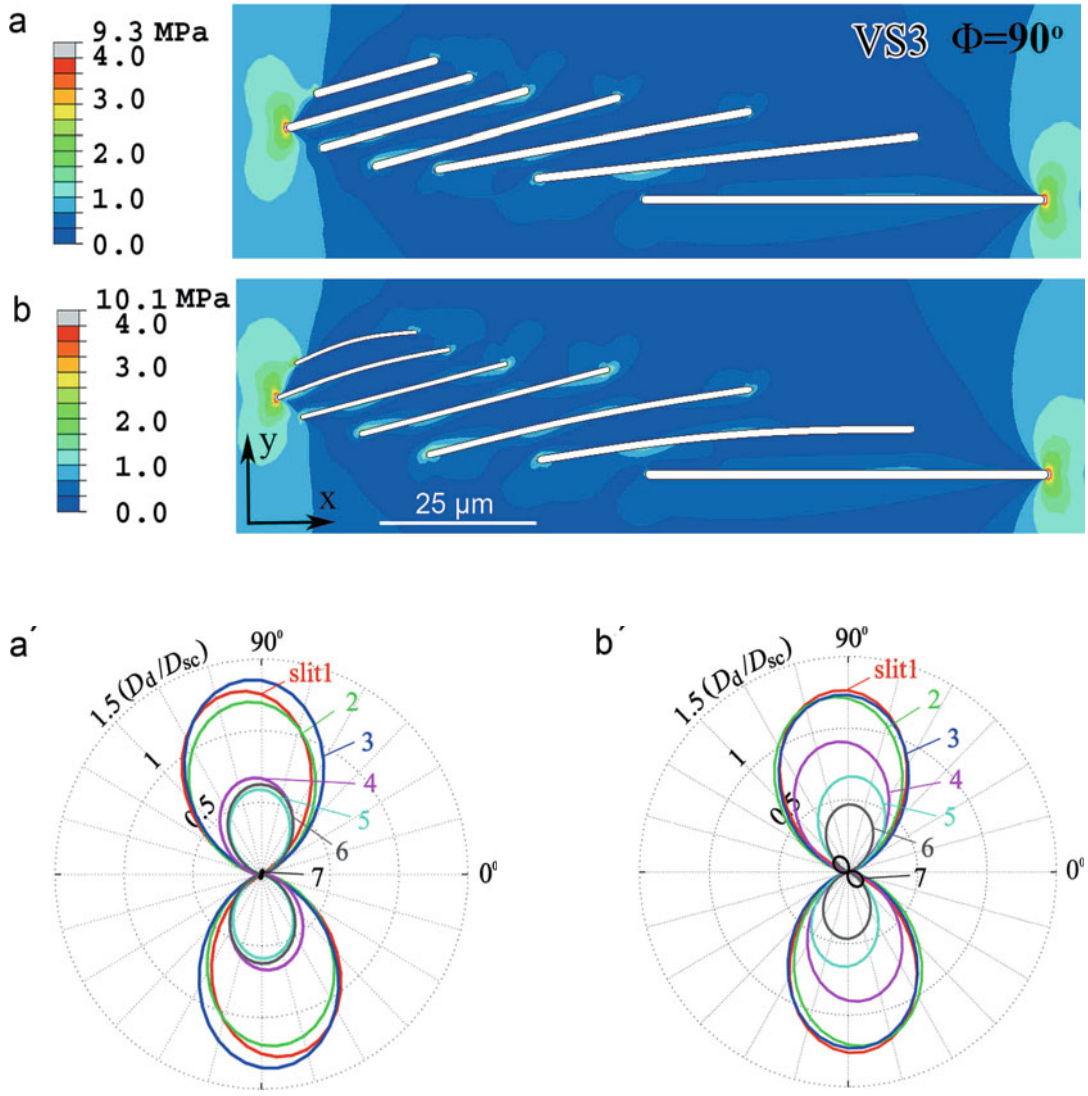


Fig. 10 *Arrays closely matching natural lyriform organs*. Finite element (FE) simulation of the directional sensitivity of seven neighboring slits as in lyriform organ VS3 (see Fig. 2f) with slit 1 being the longest and slit 7 the shortest slit in the array. **a** and **b** von Mises equivalent stresses in simple (**a**) and very detailed model (**b**) under load normal to the long axis of slit 1. **a**' and **b**' Directional mechanical sensitivity given as the ratio between slit face displacement at the dendrite's position along the slit (D_d) and the slit face displacement at mid-length of an isolated single slit (D_{sc}). Note even distribution of absolute sensitivities (range fractionation) among slits in **b**' as opposed to **a**'. (from Hößl et al. 2009, with persmission, Springer Heidelberg)

ologically (Barth and Bohnenberger 1978, Bohnenberger 1981) and the strains in the cuticle during normal behavior have been examined at the site of HS8 in life spiders using micro-strain gauges (Blickhan and Barth 1985; Brüssel 1987). In addition, the tension and pressure lines in a model tibia have been determined in early photoelastic experiments under quasi-natural load (Barth and Pickelmann 1975). Likewise the complex

deformation pattern of the elastic lines of the bars between the slits of a simplified HS8 model has been calculated from the isochromatics in an approach which treats lyriform organs as arrays of bars in a hole with the organ's outline (Blickhan et al. 1982). Importantly, spatially highly resolved white light interferometric measurements of the slit face displacements of HS8 in an intact spider, and under the same controlled load conditions as in the electrophysiological studies, are now available (Hößl et al. 2009; Schaber et al. 2011 submitted). A comparison with this data allows the reliability of the FE-models to be judged and safe quantitative predictions can be made about the slit face compressions in the original organ under biologically relevant conditions. The main conclusions are as follows.

(*i*) The slit face deformations at the location of the dendritic ending in the outer membrane predicted by the models used for FE analysis are very close (within 5 %) to the values obtained by direct measurement. This demonstrates that the deformations of slits in real organs can reliably be estimated by FE models. (*ii*) At the electrophysiologically determined stimulus thresholds which are just sufficient to elicit a nervous response of the sensory cells (Barth and Bohnenberger 1978) the slit face displacements are ≈ 1.7 nm, ≈ 14 nm, ≈ 37 nm, ≈ 43 nm, and ≈ 10 nm for slits 2,3,4,5 and 6 (slits 1 and 7 resisted proper physiological examination). The value of 1.7 nm for slit 2 is close to the value of 3 nm estimated for slit 1 when exposed to strains of –14 με which were measured with microstrain gauges during slow locomotion (Blickhan and Barth 1985). (*iii*) The sequence of the slits in the organ in regard to the magnitude of their deformation is the same in the FE simulations and in the IF measurements and follows slit length.

Conclusions and outlook

The focus of this chapter was on processes of stimulus transport and transformation. The reader interested in the nervous response of the sensory cells, their membrane physiology, transduction and encoding, is referred to reviews in Barth (2002) and French et al. (2002).

Holes like those of the arachnid slit sensilla or insect campaniform sensilla are notches in load bearing material. An engineer will try to avoid such notches because of the risk of cracks and fracture due to stress concentration which increases with the holes' aspect ratio (Peterson 1966). Obviously nature has found ways to cope with this problem by making use of the versatility of the exoskeletal cuticular material which is a fiber reinforced laminate composite (see section 3). However, the problem's solution does not emerge from the structural organization of the exoskeleton alone but from the possibility of avoiding dangerous loads. By embedding several thousands of measuring devices into the exoskeleton an arthropod like the spider *Cupiennius* will detect mechanical loads and strains with great spatial and temporal resolution. The sensory information delivered by the slits serves to avoid dangerous load situations as seen in the role slit sensilla play in synergic leg reflexes, which reinforce an imposed movement thereby pulling the leg away from large forces (Seyfarth 1978, 1985).

The sensory function of the holes rests on locally increased compliance and the local amplification of displacement. Thus measurements down to the nanometer range are possible even in a material (exocuticle) whose Young's modulus comes close to that of bone. The enormous capacity contained in the concept of embedding strain measuring sensory holes into material under load is amply demonstrated by the effect of slit structure and slit arrangement on an organ's selectivity and filtering properties. Numerous adaptations to particular measuring tasks and the fine tuning to the biologically relevant strains are possible and can actually be observed in nature. While in a single slit its length, orientation, and the position of the sensory dendrite along the slit's length are the parameters most effectively influencing mechanical sensitivity, there is a range of additional parameters to be taken into account when evaluating the

complex mechanics (and stimulus transformation) in lyriform organs. In these a much larger overall working range (re. stimulus magnitude) can be achieved by the close parallel arrangement of several differently shaped slits. The threshold stimulus amplitudes needed to elicit a nervous response were found to differ by more than 40 dB among the slits of organ HS8, and the linear parts of the slits' working ranges complement each other. Thus the linear range and the range of high increment sensitivity are greatly increased by the lyriform arrangement as compared to the single slit (Bohnenberger 1981). Generally, a single slit's deformability is reduced by the presence of closely neighbouring slits. The resulting pattern of shielding and amplifying slit face displacement can be quantitatively predicted by FE modelling. A longitudinal shift of neighbouring slits is highly influential in this regard and indeed found in most natural organs. Fan-like arrays of slits with diverging orientation make an organ sensitive to strains from a wide range of directions. Surprisingly, a large range of directional sensitivity also results from some types of parallel slit arrangements.

In addition to the versatility and refinement in the sensory periphery, differences are to be expected with regard to the central nervous convergence and integration of the signals coming from the slits of an organ. Information from slits with the same or similar response characteristics may converge to enhance the signal to noise ratio (see e. g. oblique bar model). There may be averaging by convergence of information from slits responding with different absolute sensitivity to cover a large working range (re. stimulus magnitude) (see, for example, triangular model). Similarly, the range of large difference sensitivity may be expanded by convergence of information from slits as arranged in organ HS8. These are just a few examples, which can be complemented by considering the response of an organ to stimulus direction. Clearly, with its slit sensilla the spider is able to extract the biologically relevant information from complex stimulus patterns from even the weakest signals and amidst the ever present background noise. Knowledge of the diversity of slit patterns and their mechanical properties can now also be used to predict the type of strain to be expected at the location of a particular slit arrangement.

Likewise the knowledge now available presents a challenge to engineers to design novel strain sensors based on principles found in spider slit sensilla for technical applications. The formalized quantitative nature of the predictions about sensor deformation made possible by the computational modelling approach should form a solid basis for such an endeavour.

Importantly, the tiny holes do not critically interfere with the load bearing capabilities of the material as demonstrated by their mere existence in the arthropods and also following from recent model studies on campaniform sensilla (Vincent et al. 2007). Synthetic bio-inspired slit-like sensors might serve all the technical applications strain gauges have, which still represent the most widely used devices to monitor force, motion, deformation, displacement etc. in industrial and scientific applications. These applications would include those most reminiscent of the role of the sensors in nature where they monitor movements and loads and most likely also function as sensors in feedback control loops. Such control loops are not only relevant in robotics but also for many machines and systems designed to show adaptive behavior and to react to changing working conditions (e. g. Fujimasa 1996). The bio-inspired sensors might be of particular interest for all applications where only small forces occur on small objects and where high spatial resolution is required.

A first attempt to manufacture a biomimetic strain sensor is that by Wicaksono et al. (2005) and a list of advice regarding technical sensor design based on insect campaniform sensilla can be found in Vincent et al. (2007). The goal should certainly not be to achieve an exact copy of the biological sensor but to develop micro-sensors based on the same mechanical and geometrical principles with the help of micro-technologies like silicon-micro-machining, thin film layer technologies, etching techniques and photolithography.

Acknowledgments

Much of the research reported in this chapter has received generous financial support from the Austrian Research Fund (FWF). It is my particular pleasure to also thank my many associates and students over the years, who not only substantially contributed to our research but also much enriched my life at the University.

References

Barth FG (1971) Der sensorische Apparat der Spaltsinnesorgane (*Cupiennius salei* Keys. Araneae). Z Zellforsch 112: 212–246

Barth FG (1972a) Die Physiologie der Spaltsinnesorgane. I. Modellversuche zur Rolle des cuticularen Spaltes beim Reiztransport. J Comp Physiol 78: 315–336

Barth FG (1972b) Die Physiologie der Spaltsinnesorgane. II. Funktionelle Morphologie eines Mechanoreceptors. J Comp Physiol 81: 159–186

Barth FG (1981) Strain detection in the arthropod exoskeleton. Chapter 8 in: Laverack MS, Cosens D (eds) The sense organs, pp 112–141. Blacky, Glasgow

Barth FG (1985) Slit sensilla and the measurement of cuticular strains. In: Barth FG (ed) Neurobiology of arachnids. Springer Berlin Heidelberg etc.

Barth FG (2002) A spider's world. Senses and behavior. Springer, Berlin Heidelberg

Barth FG (2004) Spider mechanoreceptors. Current Opinion Neurobiol 14: 415–422

Barth FG (2006) Spider senses – technical perfection in biology. Zoology 105: 271–285

Barth FG, Libera W (1970) Ein Atlas der Spaltsinnesorgane von *Cupiennius salei* Keys. Chelicerata (Araneae). Z Morph Tiere 68: 343–369

Barth FG, Pickelmann H-P (1975) Lyriform slit sense organs. Modeling an arthropod mechanoreceptor. J Comp Physiol 103: 39–54

Barth FG, Wadepuhl M (1975) Slit sense organs on the scorpion leg (*Androctonus australis*, L., Buthidae) J Morph 145(2): 209–227

Barth FG, Stagl J (1976) The slit sense organs of arachnids. A comparative study of their topography on the walking legs. Zoomorphology 86: 1–23

Barth FG, Bohnenberger J (1978) Lyriform slit sense organ: threshold and stimulus amplitude ranges in a multi-unit mechanoreceptor. J Comp Physiol 125: 37–43

Barth FG, Blickhan R (1984) Mechanoreception. In: Bereiter-Hahn J, Matoltsy AG, Richards R (eds). Biology of the integument. Springer, Berlin pp 554–582

Barth FG, Ficker E, Federle H-U (1984) Model studies on the mechanical significance of grouping in compound spider slit sensilla. Zoomorphology 104: 204–215

Barth FG, Nemeth SS, Friedrich OC (2004) Arthropod touch reception: structure and mechanics of the basal part of a spider tactile hair. J Comp Physiol A 190: 523–530

Bertkau Ph (1878) Versuch einer natürlichen Anordnung der Spinnen nebst Bemerkungen zu einzelnen Gattungen. Arch Naturgesch 44: 351–410

Blickhan R, Barth FG (1979) Dehnungen und Spannungen im Außenskelett von Arthropoden. GESA-Symp, Exp Spannungsanal, Braunschweig, 21 p

Blickhan R, Barth FG (1985) Strains in the exoskeleton of spiders. J Comp Physiol A 157: 115–147

Blickhan R, Barth FG, Ficker E (1982) Biomechanics in a sensory system. Strain detection in the exoskeleton of arthropods. VIIth Int Conf Exp Stress Anal, Haifa, pp.223–233

Bohnenberger J (1981) Matched transfer characteristics of single units in a compound slit sense organ. J Comp Physiol A 142: 391–401

Brüssel A (1987) Belastungen und Dehnungen im Spinnenskelett unter natürlichen Verhaltensbedingungen. Doctoral thesis, JW Goethe-Universität, Frankfurt am Main

Fratzl P, Barth FG (2009) Biomaterial systems for mechano-sensing and actuation. Nature 462: 442–448

French AS, Torkkeli PH, Seyfarth E-A (2002) From stress and strain to spikes: mechanotransduction in spider slit sensilla. J Comp Physiol A 188: 739–752

Fujimasa I (1996) Micromachines: a new era in mechanical engineering. Oxford University Press, Oxford UK

Gorbatikh L, Kachanov M (2000) A simple technique for constructing the full stress and displacement fields in elastic plates with multiple cracks. Eng Fract Mech 66: 51–63

Hößl B, Böhm HJ, Rammerstorfer FG, Müllan R, Barth FG (2006) Studying the deformation of arachnid slit sensilla with a fracture mechanical approach. J Biomech 39: 1761–1768

Hößl B, Böhm HJ, Rammerstorfer FG, Barth FG (2007) Finite element modeling of arachnid slit sensilla. I. The mechanical significance of different slit arrays. J Comp Physiol A 193: 445–459

Hößl B, Böhm HJ, Schaber CF, Rammerstorfer FG, Barth FH (2009) Finite element modeling of arachnid slit sensilla. II. Actual lyriform organs and the face deformation of the individual slits. J Comp Physiol A 195: 881–894

Kachanov M (1994) Elastic solids with many cracks and related problems. In: Hutchingson J, Wu T (eds). Adv Appl Mech 30: 259–426, Academic Press, San Diego CA,USA

McConney ME, Schaber CF, Julian MD, Barth FG, Tsukruk VV (2007) Viscoelastic nanoscale properties of cuticle contribute to the high-pass properties of spider vibration receptor (*Cupiennius salei* Keys.). J R Soc Interface 4: 1135–1143

Molina J, Schaber CF, Barth FG (2009) In search of differences between the two types of sensory cells innervating spider slit sensilla (*Cupiennius salei* Keys.) J Comp Physiol A 195: 1031–1041

Müllan R (2005) Feinbau und Rekonstruktion der Cuticulastrukturen lyraförmiger Sinnesorgane der Spinne *Cupiennius salei*. Master's thesis, University of Vienna

Peterson RE (1966) Stress concentration design factors. John Wiley & Sins Inc, New York London Sydney

Pringle JWS (1955) The function of the lyriform organs of arachnids. J Exp Biol 32: 270–278

Schaber CF, Gorb S, Barth FG (2011) Force transformation in spider strain sensors: white light interferometry (subm.)

Seyfarth E-A (1978) Lyriform slit sense organs and muscle reflexes in the spider leg. J Comp Physiol 125: 45–57

Seyfarth E-A (1985) Spider proprioception: Receptors, reflexes and control of locomotion. In: Barth FG (ed) Neurobiology of arachnids. Springer, Berlin, pp 230–248

Skordos A, Chan C, Jeronimidis G, Vincent JFV (2002) A novel strain sensor based on the campaniform sensillum of insects. Phil Trans R Soc Lond A 360: 239 -254

Vincent JFV, Clift SE, Menon C (2007) Biomimetics of campaniform sensilla: Measuring strain from the deformation of holes. J Bionic Engineering 4 (2): 63–76

Wehner R (1987) 'Matched filters' – neural models of the external world. J Comp Physiol A 161: 511–531

Wicaksono DHB, Pandraud G, Vincent JFV, French PJ (2005) Early stages of fabrication and optical characterization of new micromachined silicon strain-sensing structures inspired from the campaniform sensillum of insects. J Micromech Microeng 15: 72–81

The golden mole middle ear: a sensor for airborne and substrate-borne vibrations

18

Peter M. Narins, Urban B. Willi

Contents

Abstract

Golden moles are nocturnal, surface-foraging mammals with rudimentary vision. Several species possess massively hypertrophied mallei that presumably confer low-frequency, substrate-vibration sensitivity through inertial bone conduction. When foraging, the Namib Desert golden mole, *Eremitalpa granti namibensis*, typically moves between sand mounds topped with dune grass that contain most of the living biomass in the Namib Desert. Foraging trails are punctuated with characteristic sand disturbances in which the animal "head dips" under the sand, presumably to obtain a vibrational "fix" on the next mound to be visited. Playback experiments with substrate-borne vibrations suggest that in the absence of olfactory cues, golden moles are able to locate the food-containing mounds solely using vibrations generated by the wind blowing the dune grass. Based on middle ear anatomy, the ossicular mass distribution and the anchorage points, we have hypothesized that there are several degrees of freedom of the middle ear apparatus of *Chrysochloris asiatica*, a closely-related golden mole. A horizontal vibration, which drives the head laterally will excite the rotational mode ω_y, whereas a vertical vibration excites mode ω_z. We suggest that these two modes play the main role in inertial bone conduction in response to vibrational stimuli, since they depend on both the increased mallear mass and the displacement of its center of mass from the rotational axis. In addition to these two modes, we postulate that there is a third vibrational mode, namely the rotation of the ossicular chain about the long axis of the malleus (ω_x) in response to airborne stimuli. The transition between modes occurs between 200–300 Hz. Laser Doppler vi-

Peter M. Narins
University of California at Los Angeles
Departments of Physiological Science and
Ecology and Evolutionary Biology
621 Charles E Young Drive S
Los Angeles, CA 90095-1606, USA
e-mail: pnarins@ucla.edu

brometric measurements of the malleus head in response to vibrational stimuli in *Chrysochloris* reveal peak sensitivity to frequencies below 300 Hz. Functionally, golden moles appear to be low-frequency specialists, and it is likely that they detect prey principally through substrate conduction.

1. Introduction

Golden moles (Chrysochloridae), members of the Afrotherian clade, are small (7.6–23.5 cm) fossorial mammals endemic to sub-Saharan Africa. They exhibit a suite of specializations for life underground, are image blind (Fielden 1991) and they are "thought to rely heavily on the sense of smell and hearing" like other fossorial species (Burda et al. 1990; Francescoli 2000). The lack of light and the poor propagation of airborne sound underground limit the importance and reliability of vision and hearing in these species. The presence of conspecifics, prey and most predators is conveyed via substrate vibrations, and long-distance communication via substrate-borne signaling is considered the principal communication modality in subterranean mammals (Nevo et al. 1991).

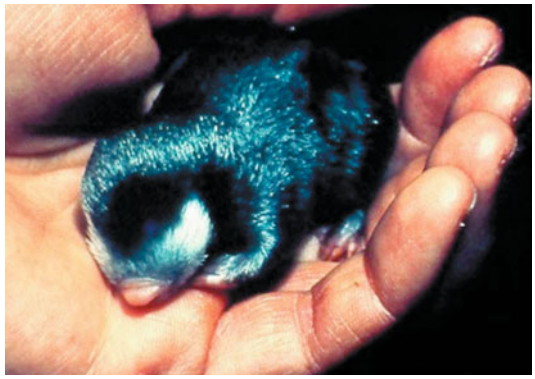


Fig. 1 The Cape golden mole, *Chysochloris asiatica* (Photo: Gary Bronner)

The proposed means by which substrate vibrations are detected in some golden mole species is inertial bone conduction. This mechanism implies the presence of a mass that impedes ossicular motion while the skull vibrates, and thereby induces relative motion between stapes and skull – an effective stimulus for the inner ear.

The Cape golden mole (*Chrysochloris asiatica*) is a member of the family Chrysochloridae, within which a variety of middle ear adaptations are found, and is the subject of our middle ear measurements (Fig. 1). Like all mammals, golden moles possess a tympanic membrane that vibrates in response to airborne sounds. Attached to its inner surface is the distalmost of the three auditory ossicles – the malleus, which in the case of the Cape golden mole, is massively hypertrophied. An overview of golden mole middle ear morphology is given by Mason (2003 a) and Fig. 2 illustrates five examples from his study.

The most obvious variations in these middle ears are the relative sizes and shapes of the mallei. Among the hypertrophied middle ear types, the malleus head shows two distinct forms: spherical or ball-shaped in *Eremitalpa* and *Chrysospalax* and club-shaped in *Chrysochloris* and *Cryptochloris*. The need for additional mass in order to detect substrate vibrations by means of inertial bone conduction seems straightforward, but what are the functional differences between the two malleus forms?

Hearing in golden moles has not been studied, but a subterranean lifestyle and poor audition are often coupled. One golden mole species that exhibits a spherical hypertrophied malleus is the Namib Desert golden mole (*E. g. namibensis*) and according to observations by Van der Vyver Nolte (1968), it shows "little to no response to sounds". This animal is sensitive to substrate

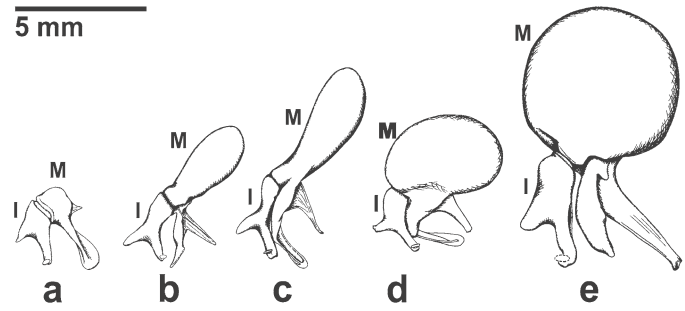


Fig. 2 Middle ear morphologies exhibited by five species of the golden mole family (Chrysochloridae). Only the malleus (M) and the incus (I) of **a** *Amblysomus hottentotus*, **b** *Chrysochloris stuhlmanni*, **c** *Chrysochloris asiatica*, **d** *Erimitalpa granti namibensis* and **e** *Chrysospalax villosus* are shown; eardrums and stapes are omitted for clarity. From Mason (2003a and 2007)

vibrations and detects and localizes dune grass-topped hummocks containing termites and other arthropods when foraging in the Namib Desert. It presumably uses a two-stage vibrational detection system such that at a distance, the animal detects the vertically-polarized Rayleigh waves emitted by the hummocks themselves in response to wind blowing the dune grass. Then when the golden mole is close to a hummock, it can sense the low-level substrate vibrations generated by the prey items as they move through the hummock (Narins et al. 1997). Moreover, these animals are capable of localizing artificial vibrational sources in the absence of olfactory cues (Lewis et al. 2006). The Cape golden mole (*C. asiatica*) exhibits a club-shaped malleus head and, perhaps not coincidentally, clear acoustic interactions between captive *C. asiatica* individuals have been reported (Willi et al. 2006b). Based on observations and on physical considerations, the two distinct morphologies of hypertrophied mallei may reflect differences in the detection abilities of these species to airborne sound.

2. Materials and methods

2.1 Animals

Cape golden moles (*Chrysochloris asiatica*) were caught in the field near Cape Town (RSA) and after at least two weeks of captivity, shipped to the University of California, Los Angeles (UCLA). Animals were housed individually, but they rarely, if ever, vocalized under these conditions. To evoke vocalizations, two animals were placed in the same tank. When they surfaced, a microphone (Brüel & Kjaer, 4134, Naerum, Denmark) placed 0.5 m above the tank recorded acoustic interactions.

2.2 Stimulation with airborne sound and substrate-borne vibrations

After surgical preparation, the caudal portion of the animal's skull was attached to an aluminum block which was fixed to a vibration exciter (Brüel & Kjaer, 4809, Naerum, Denmark). The middle ear may be stimulated acoustically with a miniature loudspeaker (MS, Knowles, CI-2960) during which the sound pressure level is recorded with a probe tube microphone (PTM, Etymotic Research, ER-7C) close to the tympanic membrane

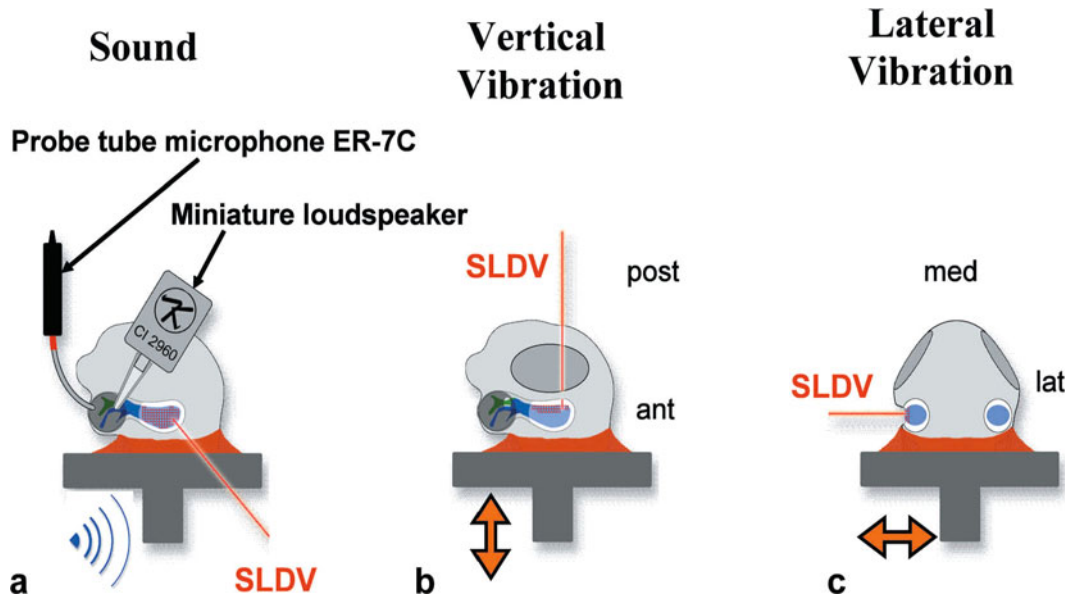


Fig. 3 **a** Caudal portion of the skull of *C. asiatica* mounted on vibration exciter. The middle ear is stimulated acoustically with a miniature loudspeaker and the sound pressure level is recorded with a probe tube microphone close to the tympanic membrane. Vibration exciter stimulating skull in **b** vertical direction, and **c** in horizontal direction. The long axis of the mallei is oriented horizontally, corresponding to that during head dipping

(Fig. 3a). Alternatively, this setup allows for vertical (Fig. 3b) or horizontal (Fig. 3c) vibrational stimulation.

2.3 Velocity measurements

The dynamic response of the malleus was measured by means of Scanning Laser Doppler Vibrometry (SLDV). The system (PSV-300, Polytec, Waldbronn, Germany) allows automated velocity measurements at multiple points (Fig. 4). It includes a built-in video camera (FCB-IX47P, Sony, Minokamo, Japan), which is optically in-line with the laser beam, and displays the selected measurement area on the computer monitor enabling grid selection, positioning and focusing of the laser beam as well as facilitating the positioning of the specimen relative to the SLDV. Such alignment is achieved by the telescopic vertical lifting column (TGC 8AWD3

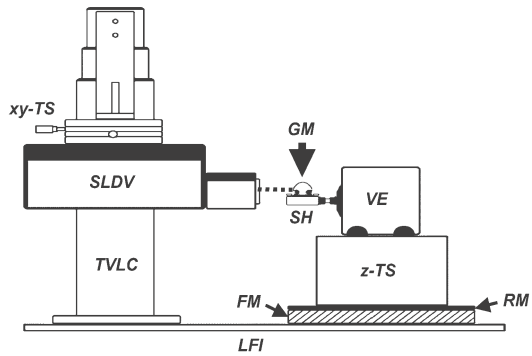


Fig. 4 Airborne/vibrational stimulation apparatus. Golden mole (GM) skull is placed on the sample holder (SH) and the beam from the Scanning Laser Doppler Vibrometer (SLDV) is focused on the head of the malleus. Beam positioning is facilitated by two translation stages (xy-TS and z-TS) and the telescopic vertical lifting column (TVLC). The skull is stimulated by the vibration exciter (VE). RM = rubber mat; FM = open cell foam mat. See text for remaining abbreviations. Modified from Willi et al. (2006b) with permission

Magnetic, Liestal, Switzerland) and an XY-translation stage (Newport, Irvine, CA, USA) in two perpendicular directions within the horizontal plane. During velocity measurements in response to both airborne and vibrational stimulation, the laser beam was in line with the medio-lateral axis of the skull, and except for the placement of the microphone and the speaker for acoustic stimulation, the two setups are identical (Fig. 4). The jack that supports the vibration exciter was placed on a rubber mat (6 mm thick) and an open-cell foam mat (25 mm thick) in order to attenuate any vibrations reaching the SLDV during vibrational stimulation. The effect of ambient vibrational disturbances was attenuated further by mounting the entire setup on a high-performance laminar-flow isolator (LFI, Newport, RS-4000, Irvine, CA, USA).

A measurement grid comprising 20–40 measurement points is superimposed on the image of the accessible distal portion of the malleus head and parts of the skull. A scan of these 20–40 points was made. Data acquisition occurred at a frequency resolution of 1 Hz for both acoustic and vibrational stimulation over the corresponding frequency band. Both amplitude and phase of the measured velocities – in case of acoustic stimulation also the amplitude of the sound pressure – were averaged five times for each measurement point.

2.4 Alignment of coordinate system

Following the recording of responses to both the airborne and vibrational stimuli, the specimen holder (aluminum block) with the skull tightly affixed to it was removed from the vibration exciter. The structures that masked the proximal portion of the ossicular chain containing the short process of the incus (SPI) and the lenticular process of the incus (LPI) during the experiment were carefully removed. All middle ear structures were preserved in order to ensure that the ossicular chain maintained its position relative to the skull. The skull was then repositioned in the measurement setup and aligned to its previous position. The built-in video camera then captured and stored an image of the entire ossicular chain.

During a laser scan, the SLDV-system assigns x- and y-coordinates to each point on the measurement grid. The position and alignment of this coordinate system was scaled and adjusted to anatomical landmarks. The origin of the coordinate system was set at the ligament of the short process of the incus (SPI), whereas the x-axis was coaxial with the long axis of the malleus and the y-axis pointed towards the lenticu-

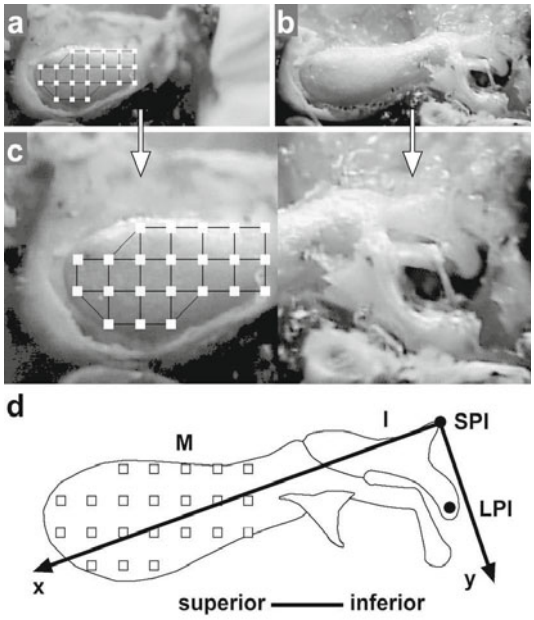


Fig. 5 Superimposition of **a** pre- and **b** post-experimental images. By superimposing the two images all structures of the ossicular chain can be referred to the coordinate system of the measurement grid **c**. **d** The coordinate system is aligned to anatomical landmarks e. g., the short process of the incus (SPI), and the coordinates of the lenticular process of the incus (LPI) were determined. Modified from Willi et al. (2006b) with permission

lar process of the incus (LPI). Superimposing the pre- and post-experimental images (Fig. 5a–c) allowed us to refer the coordinates of the measurement grid and the coordinates of the SPI and the LPI to the same coordinate system (Fig. 5d).

3. Results

We found that the long process of the incus or LPI (the attachment point to the stapes) in the golden mole middle ear responded differently to vibrational and airborne stimulation. That is, this structure's resonant frequencies for vibrational stimuli are below 200 Hz, whereas its resonant frequencies for airborne stimuli are in the range of 2 kHz.

3.1 Ossicular response

The medio-lateral motion of the LPI was calculated from the velocity measurements for the same five right ears for vibrational and airborne stimulation. The dataset was reduced to these five ears since the LPI motion was successfully reconstructed for them in response to both modalities. All responses were shifted along the logarithmic frequency axis to align the resonant frequencies at 150 Hz for substrate-borne stimulation and at 2 kHz for airborne acoustic stimulation, respectively.

For vibrational stimulation the resonant frequencies at the LPI were between 71 and 200 Hz (mean: 135.7 ± 49 Hz). Below resonance, relative motion between the skull and incus is small and the two structures move in phase; above resonance, they move out of phase by 180° (Fig. 6).

In response to airborne acoustic stimuli the LPI shows a resonance between 1.3 and 2.2 kHz (mean: 1.76 ± 0.43 kHz). At resonance, velocity amplitudes reach 1.4 ± 0.6 × 10^{-4} m/s. In four out of five specimens, a sec-

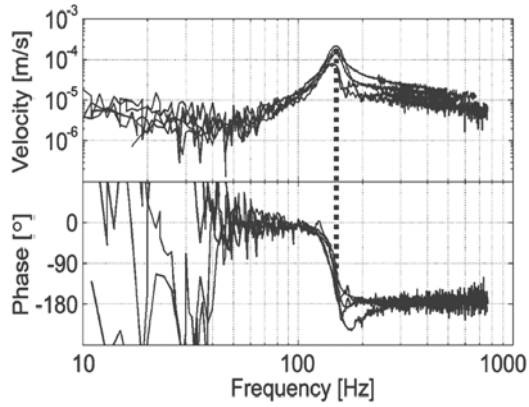


Fig. 6 Ossicular response to medio-lateral vibrational stimulation. All five responses were shifted along the logarithmic frequency scale to align all resonant frequencies at 150 Hz. Note that for frequencies below resonance, the incus and the skull are moving in phase, whereas above resonance these two structures move in antiphase. Modified from Willi et al. (2006b) with permission

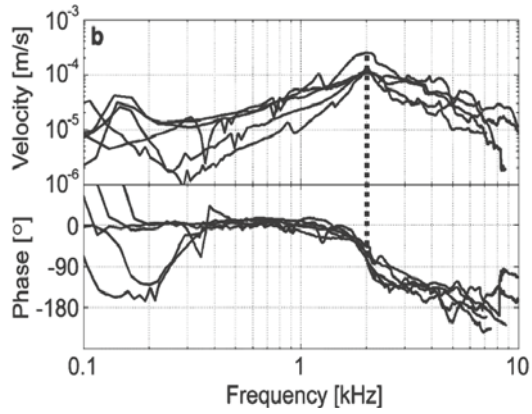


Fig. 7 Ossicular response to acoustic stimulation. All five responses were shifted along the logarithmic frequency scale to align all resonant frequencies at 2 kHz. The phase measurement shows that for frequencies below resonance, the incus and the skull are moving in phase, whereas for high frequencies (> 5 kHz), these two structures move in antiphase. Modified from Willi et al. (2006b) with permission

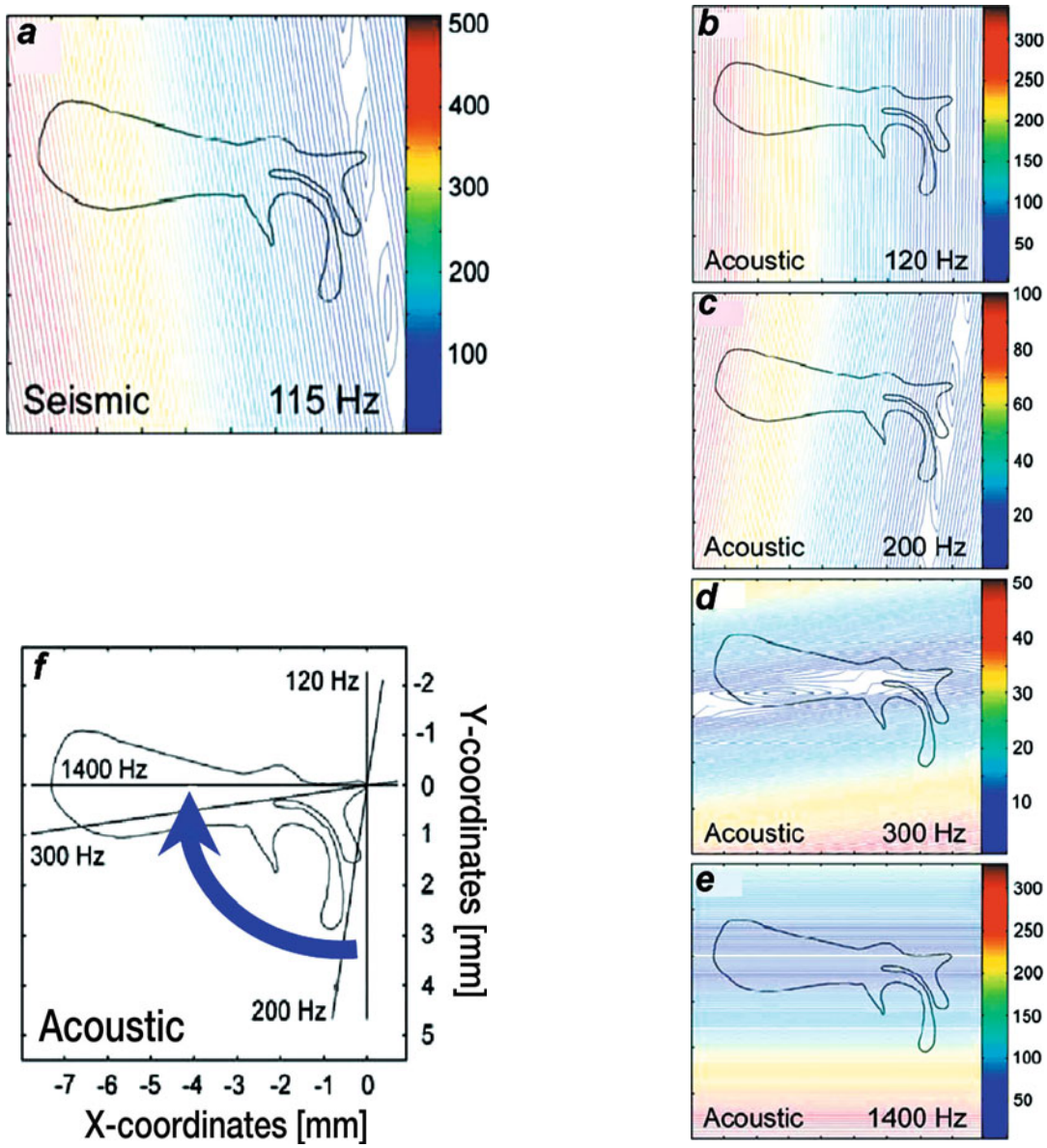


Fig. 8 Motion patterns of the malleus and incus in response to **a** vibrational and **b–e** airborne stimulation at different frequencies indicated by iso-velocity lines. The rotatary axis for each mode is defined by the iso-velocity lines of smallest amplitudes (center of blue lines). At low frequencies the rotatory axis is almost perpendicular to the long axis of the malleus **a–c**, whereas with increasing frequency it approaches the long axis of the malleus **d, e**. The transition between modes occurs between 200–300 Hz **f**. Modified from Willi et al. (2006b) with permission

ondary resonance was observed between 100 and 200 Hz (mean: 145 ± 48 Hz) which coincides well with the resonant frequency in response to the vibrational stimulation (Fig. 7).

In order to visualize the ossicular motion pattern in response to airborne and substrate-borne vibrations, we used the reconstructed iso-velocity amplitude maps generously covering the ossicular chain (Fig. 8).

The maps represent the results obtained from the left ear of one animal.

At low frequencies (115 and 120 Hz, respectively) the motion patterns for vibrational and airborne stimulation resemble each other (Fig. 8a, b). The rotatory axis passes through the SPI and is nearly perpendicular to the long axis of the malleus. For vibrational stimulation, this axis is maintained up to 600 Hz. However, with increasing frequency of the acoustic signal, this rotatory axis turns clockwise and at the resonant frequency (1.4 kHz) this axis is in line with the long axis of the malleus (Fig. 8d–e). The orientation of the rotatory axis does not change uniformly with frequency: Between 120 and 200 Hz, it rotates by about 12°, whereas between 200 and 300 Hz the axis performs a clockwise rotation of about 75°. After this rapid change it takes another 1.1 kHz in order to reach the mode at which the ossicular chain rotates precisely through the long axis of the malleus head. This mode is maintained up to about 6 kHz, above which there is a transition to the next higher mode (data not shown).

3.2 Vocalizations

Two distinct types of calls were recorded: chirps and noise bursts. Noise bursts varied in duration between 0.1–0.2 s and covered frequencies from 1–22 kHz containing most energy around 1.5 kHz (fundamental), between 4–5 kHz (third harmonic) and between 10–20 kHz (Fig. 9). The vocalization in Fig. 9 shows a short transition between a chirp (1) and a noise burst (2) but often these two types of calls are separated in time. Chirps appear as short pulses of 0.02 s but also with durations up to 0.5 s. They often contain 3–5 harmonics with most energy contained in the fundamental. High numbers of harmonics (> 10) were only observed when the chirp ended with a noise burst (Willi et al. 2006b).

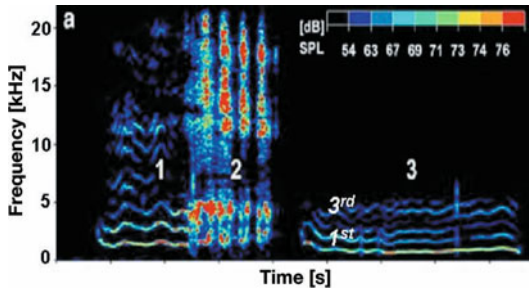


Fig. 9 Spectrogram of calls from *Chrysochloris asiatica*. During handling, the animals emitted two distinct types of sounds: chirps (call component 1 and call 3, and noise bursts (call-component 2). The two vocalizations clearly differ in the frequency band they cover and in their harmonic structure. Modified from Willi et al. (2006b) with permission

4. Discussion

That the hypertrophied malleus of the golden mole most likely comprises an adaptation for vibration detection has been discussed by several authors (Henson 1974; Kuyper 1984; Hickman 1990; Lombard and Hetherington 1993; Mason 1999; Mason and Narins 2002; Mason 2003b; Mason and Narins 2010). Moreover, the foraging behavior of the Namib Desert golden mole (*E. g. namibensis*), which also exhibits this middle ear specialization, puts this sensory modality in a meaningful ecological context (Fielden 1990; Narins et al. 1997).

4.1 Modes

The velocity profiles shown in Fig. 8 demonstrate that the ossicular chain of *C. asiatica* exhibits two distinct modes of rotation within the frequency range tested (< 10 kHz). For both vibrational modes, the malleus rotates about an axis that passes through the SPI,

and the rotatory axes of the two modes are nearly perpendicular to one another. At low frequencies (~100–200 Hz), be it in response to airborne sound or substrate-borne vibration, the club-shaped malleus head oscillates about the SPI showing maximum amplitudes at its superior end and the rotatory axis is nearly perpendicular to the long axis of the malleus (first mode). In response to higher frequency sounds (~1–6 kHz), the rotatory axis runs parallel to the long axis of the malleus head (second mode). This is not the first report of a middle ear exhibiting such distinct modes. Fleischer (1978) proposed two rotatory axes of the microtype middle ear, one in response to low and one in response to high frequencies, the two axes being perpendicular to one another. Although his ideas were critically discussed by Saunders and Summers (1982) based on velocity measurements on the mouse middle ear, Fleischer's idea and considerations clearly apply to the middle ear of *C. asiatica*.

4.2 Two vibrational modes of the golden mole malleus

The motivation for this study was the question of the functional differences between the two types of hypertrophied mallei, club-shaped and ball-shaped. We showed that the club-shaped malleus in the middle ear of *C. asiatica* responds in two distinct modes: a rotation about an axis perpendicular to the long axis of the malleus (first mode) and a rotation through the long axis of the malleus (second mode). What predictions can we make about the behavior of the ball-shaped malleus of *E. granti* or *C. villosus* (see Fig. 2)?

The hypertrophied mallei have been experimentally shown to respond to vibrational disturbances in *C. asiatica* (Willi et al. 2006a,b; Narins and Willi 2007) and there is good evidence that *E. g. namibensis* relies on the perception of substrate vibrations as well (Fielden 1990; Narins et al. 1997; Mason and Narins 2001, 2002, 2010; Mason 2007). Therefore both middle ear designs are expected to respond to vibrations. In addition, data suggest that the middle ear of *C. asiatica* is also responsive to airborne sound. What predictions can we make about the middle ear of *E. g. namibensis* with respect to airborne sound detection?

The vibrating skull is the driving force for the middle ear and the displaced center of mass of the malleus leads to stapes movements through its inertia (Fig. 10). When acoustically stimulated, the skull is not in motion and the force driving the ossicular chain is passed from the tympanic membrane to the manubrium. This results in the first vibrational mode (Fig. 10) of the malleus at low frequencies. With increasing fre-

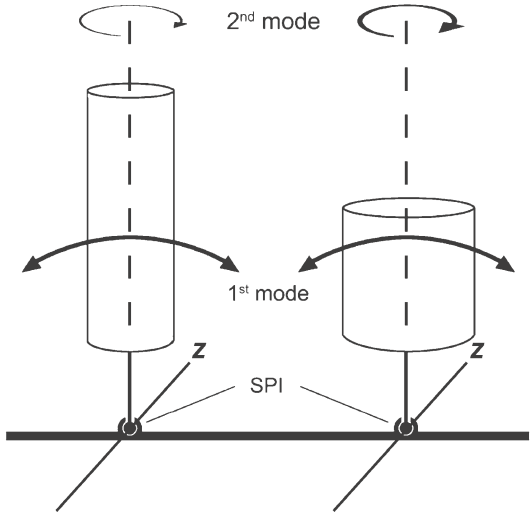


Fig. 10 Simplified schema of the two hypertrophied malleus types represented by cylinders and approximating the club-shaped (left) and ball-shaped (right) malleus for *C. asiatica* and *E. g. namibensis*, respectively. The thickness of the arrows indicates the angular moment of inertia of the ossicular chain in response to low-frequency vibrational stimulation (1st mode) and high-frequency acoustic stimulation (2nd mode). SPI: tip of the short process of the incus. Modified from Willi et al. (2006b) with permission

quency, though, the mass of the malleus increasingly impedes this mode.

In order to partially overcome the impedance of the mass, the middle ear design allows the ossicles to respond in the second mode, which has its rotatory axis passing through the center of mass. This is possible for both middle ear designs, the club-shaped and ball-shaped malleus. However, in the second mode the two designs differ significantly in their angular moment of inertia defined by:

$$I_2 = {}^1/_2 \, m \, r^2 \qquad (1)$$

where I_2 is the angular moment of inertia of the second mode (Fig. 10), m is the mass of the cylinder and r its radius. The thickness of the arrows in Fig. 10 indicates the relative magnitude of the angular moment of inertia. If the angular moment of inertia for the first mode is identical in both designs, then that for the second mode is necessarily different, being larger for the ball-shaped malleus (Fig. 10, right model) compared to the club-shaped malleus (Fig. 10, left model). As a result the responsiveness of the club-shaped malleus would exceed that of the ball-shaped malleus. When the radius is doubled, the angular moment of inertia is quadrupled. An extreme version of a large malleal radius is found in the middle ear of *Chrysospalax villosus* (Fig. 2 e) but in *E. g. namibensis* it is also quite large compared to that of *C. asiatica*.

A ball-shaped design is volume efficient, implementing a vibration detector occupying minimal space. In order to allow the middle ear to respond to higher frequencies with a second mode, the radius of the malleus head is reduced. In *C. asiatica* the radius of the malleus is reduced but it also increases towards the distal end, which shifts the center of mass further away from the anchorage point (SPI). The further away the center of mass is shifted in the club-shaped malleus, the thinner the cylinder may become, while still maintaining the same angular moment of inertia as the ball-shaped design, since the angular moment of inertia for the first mode is determined by:

$$I_1 = {}^1/_3 \, m \, l^2 \qquad (2)$$

where I_1 is the angular moment of inertia of the first mode (Fig. 10), m is the mass and l is the distance of the center of mass from the suspension point (highly simplified model).

The unusual malleus morphology in the middle ear of *C. asiatica* has been shown to respond to substrate vibrations and airborne sound. Its design appears to be an elegant solution to the problem of how a middle ear can exploit increased ossicular inertia for the detection of low-frequency substrate vibrations at little cost to the animal's high-frequency airborne sound-detection capability.

Clearly, there are many fascinating questions that remain unanswered: Do the two vibrational modes of the golden mole malleus translate to disparate mechanical input to the inner ear via the stapes? Can the inner ear differentiate between these two separate inputs? How do the viscoelastic and other properties of the malleus confer frequency dependence on the middle ear? How do golden moles localize a vibrational source? This question is especially interesting since in *E. g. namibensis*, the mallei of the two ears are not parallel. This raises the exciting possibility of azimuth-dependent differential mallear motion, a unique adaptation for vibrational localization in vertebrates. Experiments are currently planned to test this intriguing hypothesis.

Acknowledgments

Thanks to Gary Bronner for supplying us with Cape golden moles for the middle ear measurements. We also thank the staff of the Division of Laboratory Animal Medicine at UCLA, particularly Dr. Marcelo Couto and Dr. Greg Lawson for their helpful collaboration. This research was supported by NIH grant no. DC00222 to PMN. All animal procedures were approved by the Animal Research Committee at the University of California, Los Angeles, USA.

References

Burda H, Bruns V, Müller M (1990) Sensory adaptation in subterranean mammals. In: Nevo E, Reig OA (eds) Evolution of subterranean mammals at the organismal and molecular levels. Wiley-Liss, New York, pp 269–293

Fielden LJ (1991) Home range and movements of the Namib Desert golden mole, *Eremitalpa granti namibensis* (Chrysochloridae). J Zool London 223: 675–686

Fleischer G (1978) Evolutionary principles of the mammalian middle ear. Adv Ant Embryol Cell Biol 55: 1–70

Francescoli G (2000) Sensory capabilities and communication in subterranean rodents. In: Lacey EA, Patton JL, Cameron GN (eds) Life underground: The biology of subterranean rodents. The University of Chicago Press, Chicago, pp 111–144

Henson OW (1974) Comparative anatomy of the middle ear. In: Keidel WD, Neff WD (eds) Handbook of sensory physiology V/1, Auditory system. Springer, Berlin, pp 39–110

Hickman GC (1990) The Chrysochloridae: Studies toward a broader perspective of adaptation in subterranean mammals. In: Nevo E, Reig OA (eds) Evolution of subterranean mammals at the organismal and molecular levels. Wiley-Liss, New-York, pp 23–48

Kuyper MA (1984) Golden moles. In: MacDonald D (ed) The encyclopedia of mammals V/2. Allen & Unwin, London, pp 764–765

Lewis ER, Narins PM, Jarvis JUM, Bronner G, Mason MJ (2006) Preliminary evidence for the use of microseismic cues for navigation by the Namib golden mole. J Acoust Soc Am 119: 260–1268

Lombard RE, Hetherington TE (1993) Structural basis of hearing and sound transmission. In: Hanken J, Hall BK (eds) The skull V/3. University of Chicago Press, Chicago, pp 241–302

Mason MJ (1999) The functional anatomy of the middle ear of mammals, with an emphasis on fossorial forms. Unpublished doctoral thesis, Cambridge University Press, Cambridge, UK

Mason MJ (2003a) Morphology of the middle ear of golden moles (Chrysochloridae). J Zool 260: 391–403

Mason MJ (2003b) Bone conduction and seismic sensitivity in golden moles (Chrysochloridae). J Zool 260: 405–413

Mason MJ (2007). Massive mallei in moles: Middle ear adaptations subserving seismic sensitivity. Proc Inst Acoust 29: 69–76

Mason MJ, Narins PM (2001) Seismic signal use by fossorial mammals. Amer Zool 41: 1171–1184

Mason MJ, Narins PM (2002) Seismic sensitivity in the desert golden mole (*Eremitalpa granti*). A review. J Comp Physiol A 116: 158–163

Mason MJ, Narins PM (2010) Seismic sensitivity and communication in subterranean mammals. In: O'Connell-Rodwell CE (ed) The use of vibrations in communication: Properties, mechanisms and function across taxa. Kerala: Transworld Research Network 121–139

Narins PM, Lewis ER, Jarvis JJUM, O'Riain J (1997) The use of seismic signals by fossorial southern African mammals: A neuroethological gold mine. Brain Res Bull 44: 641–646

Narins PM, Willi UB (2007) Mostly malleus: Ground sound detection by the remarkable ear of the golden mole. Proc 19th Internatl Congress on Acoustics, Madrid, Spain, BIO-02-003: 1–5

Nevo E, Heth G, Pratt H (1991) Seismic communication in a blind subterranean mammal: a major somatosensory mechanism in adaptive evolution underground. Proc Nat Acad Sci USA 88: 1256–1260

Saunders JC, Sommers RM (1982) Auditory structure and function in the mouse middle ear: an evaluation by SEM and capacitive probe. J Comp Physiol A 146: 517–525

Van der Vyver Nolte H (1968) The external morphology and functional anatomy of the cranial region in the Namib golden mole *Eremitalpa granti namibensis* Bauer und Niethammer, 1959. M Sc Thesis. University of Pretoria, Pretoria

Willi UB (2003) Middle ear mechanics: The dynamic behavior of the incudo-malleolar joint and its role during the transmission of sound. Un-

published doctoral thesis. University of Zürich, Zürich

Willi UB, Bronner GN, Narins PM (2006a) Middle ear dynamics in response to seismic stimuli in the Cape golden mole (*Chrysochloris asiatica*). J Exp Biol 209: 302–313

Willi UB, Bronner GN and Narins PM (2006b) Ossicular differentiation of airborne and seismic stimuli in the Cape golden mole *(Chrysochloris asiatica)*. J Comp Physiol A 192: 267–277

Insect inertial measurement units: gyroscopic sensing of body rotation

19

Thomas Daniel, Zane Aldworth, Armin Hinterwirth, Jessica Fox

Contents

Abstract

All flying insects require both visual and mechanosensory information to effectively control flight in complex environments. Sometimes, however, visual systems may not have sufficiently short response times to accommodate for rapid perturbations in the flight path. Thus insects also rely on fast and precise mechanosensory systems as part of their flight control mechanism. We review a subclass of mechanosensory systems that serve as gyroscopic organs – otherwise known as inertial measurement units in engineering systems. We review current neurophysiological and behavioral data for two putative biological gyroscopes: cranefly halteres and the antennae of hawkmoths. In addition we present data that further clarifies the gyroscopic function of moth antennae, showing that they directly measure pitching motions. We also introduce new data on mechanosensory systems in the wings of moths and suggest that these too could provide critical inertial sensory information to flying animals.

Thomas Daniel
University of Washington, Department of Biology
Seattle, WA 98195-1800, USA
e-mail: danielt@u.washington.edu

1. Introduction

Flying insects acquire and process information about their environment to control their movement in a complex and dynamic world. They do so over a vast range of temporal and spatial scales, from long distance navigation occurring over thousands of kilometers (*e.g.* monarch butterflies: Etheredge et al. 1999), to landmark navigation for daily foraging (Srinivasan et al. 1996) to shorter term visual information processing associated with pollination behavior (Srinivasan 2006, Sprayberry 2009), predator-prey interactions (Lehmann 2009) and mating behaviors (Land and Collett 1974).

Of all insect sensory modalities associated with flight control, mechanosensory systems operate with the fastest time scales, sometimes as short as a few milliseconds between stimulus onset and behavioral response. Some of these mechanosensory processes are associated with proprioreceptors

tive inputs critical to flight control such as head stabilization for visual signal processing (Collett and Land 1975; Sandeman 1980; Huston and Krapp 2009), and wing position detection (Frye 2001). Others are involved in exteroreceptive information processing including airflow detection (Niehaus 1981; Sane and McHenry 2009; Fratzl and Barth 2009), gravitational sensing (Kamikouchi et al. 2009) and, of great importance to flight control, gyroscopic sensing via specialized inertial measurement organs (Frye 2007; Daniel et al. 2008; Sane and McHenry 2009)

Since the times needed to process visual information can be quite large relative to the time-scales of movement control, particularly under low light conditions (Warzecha and Egelhaaf 2000; Theobald et al. 2009), alternate sensory modalities are critical for rapid course correction. As such, animals rely on multimodal information processing to coordinate movement over a broad range of time scales, using mechanosensory reflexes for responses to rapid perturbations and vision for slower processes. Here we review recent research and controversies surrounding mechanosensory systems associated with inertial measurement in insect flight control.

2. Halteres and antennae act as vibrating structure gyroscopes

Vibrating structure gyroscopes (VSGs) are commercially available inertial measurement sensors whose inspiration may have followed from our understanding of specialized sensory structures in flying insects. The principles underlying these sensors are similar those for rotating gyroscopes: a proof mass moving in an arc subject to rotation in an orthogonal plane experiences an acceleration a that is proportional to the cross product of the two rotational velocities v and Ω: $a = -2(v \otimes \Omega)$. The torque (τ) associated with that acceleration is proportional to the product of the mass m and that acceleration: $\tau = -2m(v \otimes \Omega)$. For vibrating systems this torque will have components at the driving frequency and twice that frequency.

2.1 Halteres

The most well known biological structure that operates as a VSG is the haltere, found in all true flies, including fruit flies, house flies, and their relatives, mosquitoes and crane flies (Pringle 1948). These organs are derived from hind-wings that are reduced to club-shaped structures (Fig. 1), and are oscillated in a plane perpendicular to the long axis of the body (Nalbach 1993). Thus the angular velocity (v) in the acceleration equation above is a periodic function of time. Body rotations in axes orthogonal to the haltere stroke plane induce Coriolis forces, and these forces yield strains on the cuticle of the haltere at the driving frequency and double that frequency. Each haltere contains at its base fields of sensory structures (campaniform sensilla) which have been shown to be directly connected via a single neural synapse to the muscles controlling flight (Fayyazuddin and Dickinson 1996). Recent theoretical work (Thompson et al. 2009) has shown that the movement of halteres subject to Coriolis forces takes on complex lemniscate tip trajectories that could be encoded to produce an effective flight control system. They suggest that if the neural system measured mid-stroke motions with great precision, then one could construct a biologically inspired inertial rate sensor.

Despite overwhelming evidence suggesting that halteres respond to these small

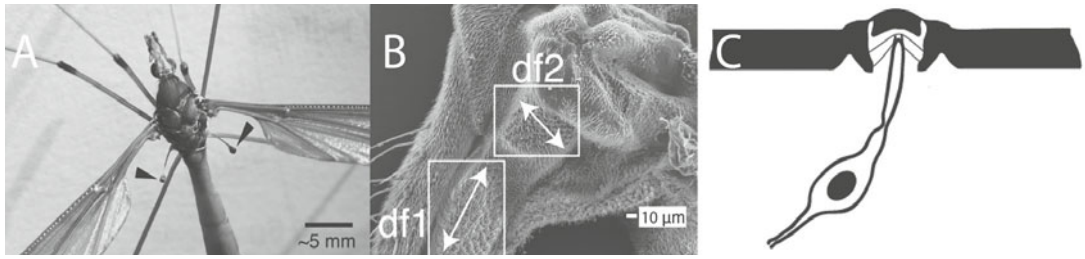


Fig. 1 The crane fly *Holurusia* sp (panel **A**) shows the location of halteres (arrows). The base of the haltere has several fields of campaniform sensilla (panel **B** shows the df1 and df2 fields). Each tiny dome in the fields is a single sensillum shown diagrammatically in panel **C** (after Kiel 1997)

Coriolis forces, there was until quite recently no strong evidence indicating that haltere campaniform sensilla could provide the appropriate temporal resolution and motion phase locking. We (Fox and Daniel 2008) asked if indeed the single neurons associated with these strain sensors could provide that information to the nervous system and its flight control centers. Using an extraordinarily large Dipteran species, the cranefly *Holorusia* (Fig. 1), we were able to mechanically stimulate the haltere and record intracellularly from the neurons receiving strain information.

Neurophysiological data from those intracellular recordings of haltere sensory neurons showed that they are able to provide the appropriate temporal and spatial strain resolution to serve a gyroscopic sensory function (Fox and Daniel 2008). The latency from application of a punctate stimulus at the distal tip of the haltere to action potential detection in the sensory cell's axon midway between sensilla and the pterothoracic ganglion was measured to be about 4 ms across a population of 19 different cells. This short response delay, coupled with the known single-synaptic geometry of the haltere-to-flight muscle circuit (Fayyazuddin and Dickinson 1996), indicates that processing for mechanosensory input from the halteres will always be far faster than the equivalent processing speeds of the multi-synaptic pathway from retina to flight control centers (at least 30 ms, Collett and Land 1975). Additionally, haltere sensory neurons are able to encode oscillatory motions at frequencies well above those critical for Coriolis force measurement (Fox and Daniel 2008).

These gyroscopic sensors phase-lock with a level of precision rarely seen for any biological sensory structure. Phase locking, as measured by vector strength (Fox and Daniel 2008), may reach a theoretical maximum of 1.0. Barn owl auditory neurons, for example, are associated with exquisite sound localization and show a vector strength that lies between 0.9 and 0.95 (Köppl 1997). Haltere neurons, on the other hand, have a vector strength of approximately 0.98 – the highest reported thus far. If phase-locked strain sensing is crucial for gyroscopic force measurement, then halteres are well equipped to provide that sensory function.

Finally, there remain several tantalizing issues associated with biological gyroscopes. First, they are able to respond to remarkably small motions. A few degrees of tilt are sufficient to cause depolarization of the sensory neuron. Additionally, the structures that contain the dendrites (campaniform sensilla, Fig. 1C) of these sensory neurons are themselves highly asymmetric, suggesting that they filter motion in a single plane. Thus, while the haltere may experience complex three-dimensional patterns of strain, it is possible that the campaniform sensilla to which the sensory cell dendrites

attach could provide a directional filter for information encoding.

2.2 Moth antennae

Like flies, moths and butterflies use both visual and mechanosensory inputs in flight control. The large delays involved in visual processing imply that vision alone is not sufficient to provide all the necessary information required for flight control (Hedrick and Daniel 2006). The relatively quick processing of information generated by mechanosensory structures could potentially supplement the visual information, much like in case of the halteres (Dickinson 1999). Until quite recently, however, the role of mechanosensory input remained poorly explored for Lepidoptera. Frye (2001), for example, showed that a single sensory neuron comprising the stretch receptor in the hind-wing of the sphinx moth *Manduca sexta* is crucial for flight control. It encodes at critical frequencies and its ablation compromises the visually-mediated flight response. However, it has not been shown that this cell is critical for balance control. Finally, wind sensing has been suggested as one of the key mechanosensory functions of antennae in Lepidoptera (Niehaus and Gewecke 1981, Niehaus 1981). Such a role of antennae for regulating flight speed has been demonstrated in a variety of insects, such as bees (Heran 1959), flies (Burkhardt and Schneider 1957), or locusts (Gewecke 1970). What remains unclear is whether antennae subserve an additional function in balance stabilization, as organs that can detect effects of small-scale turbulence on an animal's flight path in order to mediate rapid corrective maneuvers.

Sane et al. (2007) recently suggested that the antennae of moths serve a gyroscopic function. Indeed, they satisfy all of the critical criteria we established above for sensory gyroscopes. Together, the evidence supporting this role is quite strong. For example, antennal removal significantly compromises flight; subsequent reattachment of antennae restores flight. Additionally, the base of the antenna contains both the Böhm's bristles and Johnston's organ, specialized mechanoreceptors that send neural projections to the motor control centers of the central nervous system (Kloppenburg et al. 1997).

Like the halteres of flies, the strain sensors at the base of antennae respond with very low latencies (about 2 ms to the motor control center of the brain: Sane et al. 2007). Also like halteres, they can encode at the critical frequencies associated with Coriolis forces. In flying moths, for example, the wingbeat frequency is approximately 25 Hz suggesting that strain signals will occur at 25 and 50 Hz. Unlike haltere neurons which show broad tuning (and linear encoding) over a very broad frequency range, the neural response of antennal mechanosensors suggests that they are tuned to 50 Hz, the frequency most uniquely associated with Coriolis forces (Sane et al. 2007).

There is still some controversy about whether antennal mechanosensory systems measure inertial forces associated with rotation, forces associated with aerodynamic drag (Niehaus 1981) or even gravitational forces for posture control (Kamikouchi et al. 2009). Indeed, it is entirely possible that they serve all three roles. But the purported gyroscopic function for antennae, despite the neurophysiological evidence and observations from animals with antennae removed (Sane et al. 2007), is still an open question. Those data alone do not preclude wind sensing as the sole mechanosensory function for antennae: antennectomy leads to loss of wind detection as well as loss of putative Coriolis force detection in freely flying animals. To demonstrate rotational sensing, it is critical to separate visual, wind and rotational stimuli. We have done so using a flight simulator modeled after those

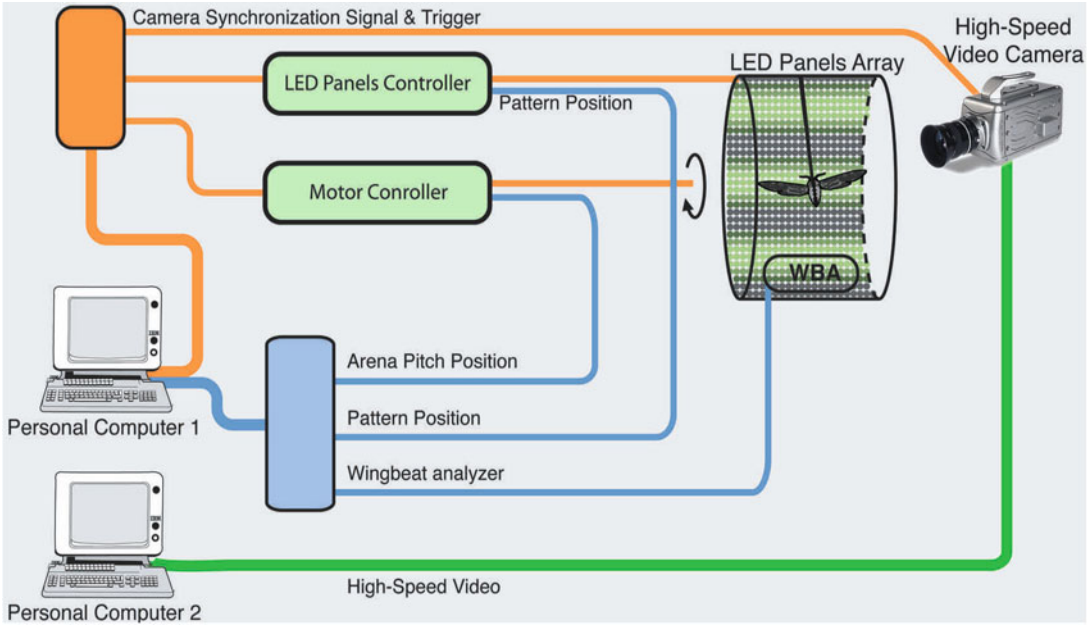


Fig. 2 The flight simulator used in studies of antennal mechanosensory function. One computer controls both the visual world, consisting of a cylindrical arena lined with LED panels. In addition that computer also controls a stepper motor that can independently rotate the arena with the tethered moth and synchronizes these with a high speed digital video camera (Phantom v5 camera). A second computer records the data from the high speed digital video camera to document wing and abdominal motions that occur in response to visual or mechanical pitch stimuli (or a combination of the two)

developed by Reiser and Dickinson (2008) in which visual and mechanical rotational stimuli can be delivered independently to tethered animals (Fig. 2).

Results from the flight simulator show that visual and mechanical stimuli drive a combination of reflexive maneuvers in flying *Manduca (*sphinx moth*)*. In addition to visually stimulated head and wing motions, there are, as in most large flying insects, pronounced reflexive abdominal motions. Thus visual pitch stimuli lead to dorso-ventral flexions of the abdomen as well as changes in the stroke amplitude of the wing elevation angle (Hinterwirth and Daniel 2010). Similarly, visual yaw stimulation leads to lateral abdominal motions. Interestingly, animals exposed to periodic mechanical pitch stimuli with no changes in visual or wind stimuli also show phase locked abdominal flexions and extensions. It is clear, therefore, that some sensory modality is capable of detecting body rotations.

A simple experimental manipulation, inspired by those of Sane et al. (2007), can resolve the issue of Coriolis force detection by antennae. If antennae are involved, in part, for reflexive control of the abdomen in response to pure rotational stimulation, then removal of the antennal flagellum (the proof mass for detecting Coriolis forces) would eliminate that response. Moreover, since the critical mechanosensory cells lie proximal to the antennal flagellum, the neural machinery associated with Coriolis force detection remains intact following antennectomy. Hence, reattaching antennae would recover the mechanically stimulated abdominal reflexes.

Our data show that animals subject to a 3 Hz mechanical pitch stimulation (with no changes in either visual or wind stimulation)

respond with phase locked abdominal flexions of about 5 degrees (Fig. 3). And, as predicted, removal of the flagellum eliminates that response and subsequent antennal reattachment recovers the response.

The original abdominal reflex is not significantly different from that associated with animals having antennal reattachment (Fig. 4). These data, therefore, provide unequivocal evidence that antennae are indeed detecting rotations of the body. Indeed for any animal with an appendage that vibrates actively or passively, a Coriolis force will occur whenever the animal turns during locomotion. If, in addition, that appendage is equipped with motion (or strain) sensitive sensory cells, it could detect that force. The challenge remains, however, to determine whether the nervous system uses such information in locomotory control.

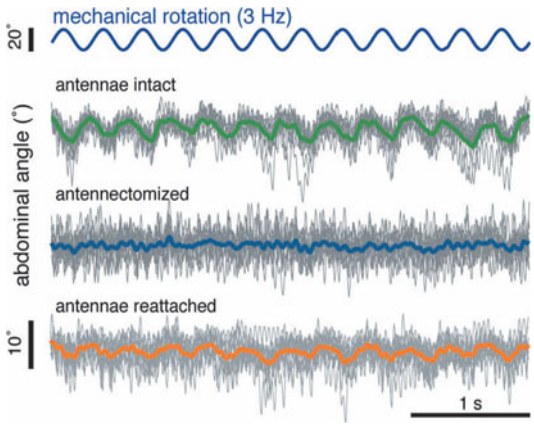


Fig. 3 The motion of the arena (blue trace) for a periodic mechanical pitching signal is plotted against time. For one individual we also plotted the abdominal flexion angle that occurs in response to that pitch stimulus for the intact animal (*Manduca sexta*) (12 trials shown in grey, mean shown in green), with its antennae removed (12 trials shown in grey, mean shown in blue) and with its antennae reattached (12 trials shown in grey, mean shown in orange). There was no visual stimulus provided to this animal

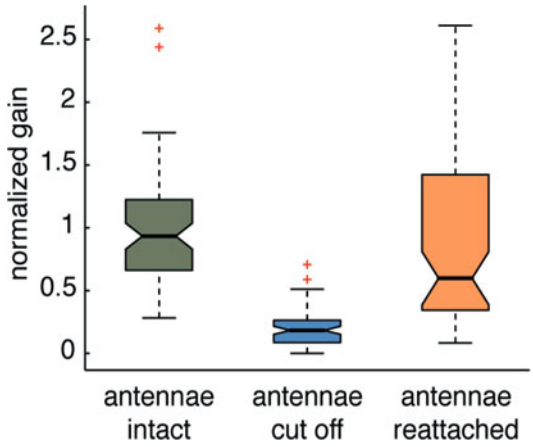


Fig. 4 The normalized gain of the abdominal flexion of *Manduca sexta* is shown for the three treatments summarized in Fig. 3: intact animals, those with antennae removed, those with antennae reattached). The response gain is the abdominal flexion amplitude divided by the amplitude of the pitch stimulus. It is normalized to the mean gain of intact animals. We analyzed 7 animals, each with 6 trials. Antennectomized animals are significantly different (t-test, $p < 0.01$) from intact animals. The gain response of those with reattached antennae are not significantly different from intact animals

3. Might wings or other structures provide data about inertial forces?

The examples above set the stage for understanding the more complex and enigmatic roles of wing strain sensors. While all insect wings contain campaniform sensilla, Lepidoptera have particularly large arrays of them. In *Manduca* there is a rich set of campaniform sensilla distributed over both the upper and lower surface of the wings (Fig. 5). Approximately 300 sensors adorn the dorsal and ventral surfaces of the front wing, with the majority located in fields of sensilla at the base. Approximately 130 more sensors are distributed over the surfaces of the hind wing, again with the bulk of them located

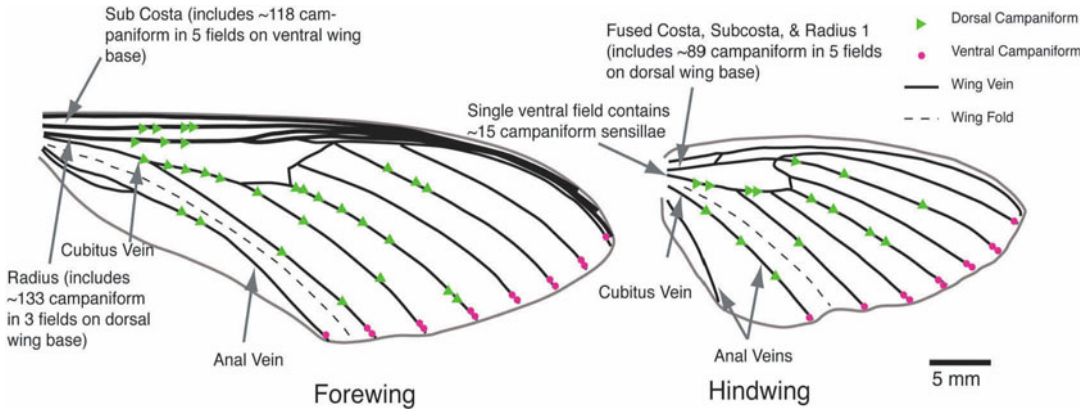


Fig. 5 Diagrammatic representation of the strain sensor distribution on the fore- (left) and hind-wing (right) of *Manduca sexta*. Locations of individual sensors on the dorsal (triangles) and ventral (circles) wing surfaces, as well as the approximate location of the fields of campaniform sensilla at the wing base are shown. Adapted from Dombrowski (1991)

near the wing base (Dombrowski 1991). In general we are unsure about how sensory information from these structures is used in any flying insect. We have, therefore, a reverse engineering problem: identify what sensory information from wing campaniforms is used in flight control.

We use evolutionary history to inform us about one potential role for these sensors in the wings. As mentioned above, dipteran halteres are derived from the hind wings of some ancestral progenitor. Similarly, the forewings of strepsipteran insects have been modified through evolutionary processes to serve the same role as halteres in the diptera (Pix et al. 1993). For natural selection to favor a gyroscopic function for these structures, the wings themselves might have provided at least rudimentary Coriolis force information to the antecedents of Diptera and Strepsiptera. Indeed, since wings are essentially proof masses oscillating in a plane, we can assert that they must experience a Coriolis force whenever the animal's body rotates in a plane orthogonal to that of the wing stroke. However, while we can reasonably posit the existence of such a force, we do not know whether it is encoded and processed by the central nervous system. If wings not only drive flight, but also detect inertial moments, then they would be in the unique position of simultaneously serving as organs of both sensation and actuation.

We turn to aerodynamic considerations to identify a second potential use of the campaniform sensilla. We know that many large insect wings deform quite dramatically during flight (Combes and Daniel 2003), and that such deformation is dominated less by aerodynamic loading than by inertial and elastic processes for wings having structural properties similar to insect cuticle (Daniel and Combes 2002). Hence high wing beat frequencies drive inertial bending of the wings, which has important aerodynamic consequences. For example, by applying high-speed digital particle image velocimetry to isolated and robotically actuated wings we were able to show that wings subject to large amounts of bending (those from recently emerged moths) generate significantly more lift-favorable momentum flux than stiffer wings of the same mass (Mountcastle and Daniel 2009). The mechanisms that underlie this increased momentum flux are still unclear but likely relate to favora-

ble passive wing pitching motions as well as elastic bending waves that propagate in a chordwise direction from the leading to trailing edge of the wing.

Taken together, the information we have above suggests two possible functions for information encoded within the wing campaniform sensilla. The first role is gyroscopic sensation analogous to the halteres of the Diptera and Strepsiptera, for which the wings themselves are the evolutionary precursors. In this capacity we would expect that the campaniform sensilla would be particularly good at representing forces occurring at the Coriolis frequencies (25 and 50 Hz in *Manduca*, corresponding to the wingbeat and its first harmonic, respectively). The second role is in encoding aerodynamic forces on the wing, suggested by the fact that wing bending occurs and is aerodynamically important. If the campaniform sensilla operated in this mode we would expect to see sensitivity to the relatively higher frequency signals that would occur due to propagation of bending waves across the surface during each wing stroke. In either case we would like to know the precision with which these sensors could encode information about the relevant signal, as well as how much information they encode.

Drawing on our experience with halteres and antennae, we examined the encoding properties of these sensory structures, asking at what frequencies they could encode information and with what level of precision. We used identical techniques as above (intracellular recording from the sensory axons, mechanical stimulation of the wings)

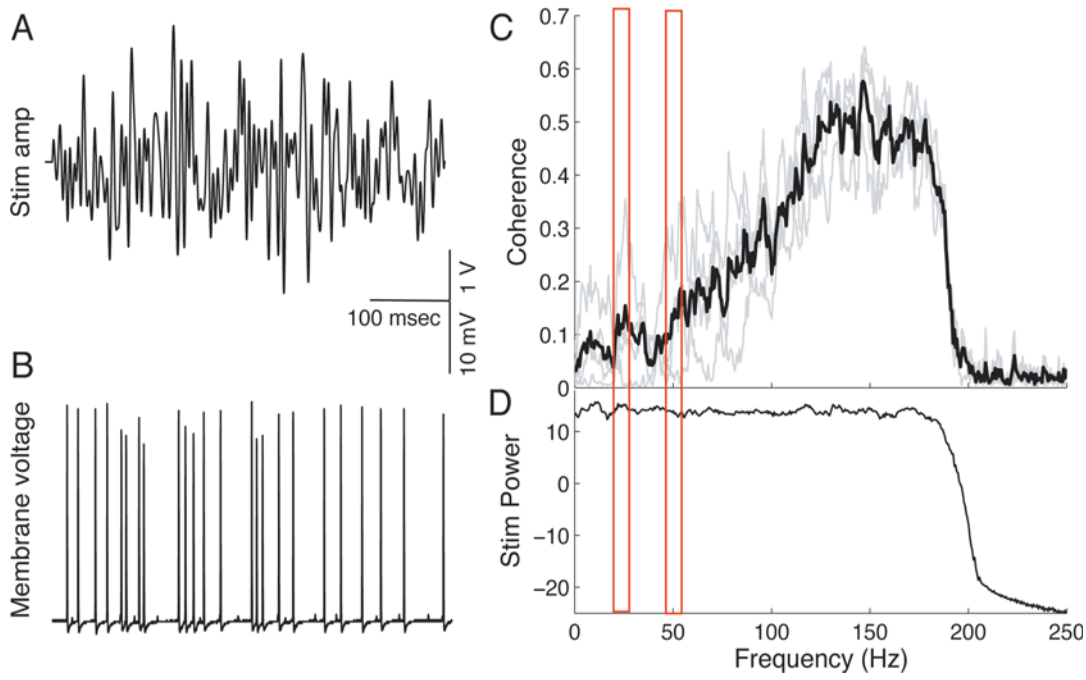


Fig. 6 Electrophysiological data from sensory cells supplying wing campaniform sensilla. Panel **A** shows a single trace of the band-limited mechanical white noise stimulus with the corresponding intracellular neuronal response to that stimulus in panel **B**. Panel **C** shows the coherence curves for four individual cells (light gray traces)and the population mean (heavy black). Panel **D** shows the power spectrum of the corresponding stimulus. The frequency regions corresponding to Coriolis forces are indicated by red boxes

to characterize their response functions. Repeated segments of 2–200 Hz band-limited white noise were provided during stimulation, allowing us to calculate temporal precision using the jitter metric (Berry et al. 1997), frequency tuning from the coherence function (Theunissen et al. 1996), and a lower limit to the information rate using the stimulus reconstruction technique (Rieke et al. 1997). Figure 6 shows the results of these analyses performed on data from the afferent neurons of four different wing campaniform sensilla. For these four cells the precision of spiking across multiple repeats of a single stimulus was 0.29 ± 0.28 msec (mean ± SD), coherence was significant from 2–200 Hz (the entire range of our band limited stimulus) with a peak at ~150 Hz, and the information about the stimulus contained within the responses was 103.2 ± 22.5 bits/sec (mean ± SD), corresponding to 2.5 bits/action potential. These information transmission rates compare favorably with the highest such estimates yet reported (Borst and Theunissen 1999).

The broad frequency tuning of these strain sensors does not allow us to determine whether wing campaniform sensilla act purely as Coriolis sensors or as sensors of aerodynamic forces. In fact, the high precision information encoding rates of their sensory cells suggest that they could serve multiple roles. In addition to potentially encoding Coriolis forces and providing information about dynamic bending waves, they could also provide data on aerodynamic flutter and other high frequency events.

Conclusions

Insects are imbued with a rich set of strain sensors, all of which provide information about bending and deflection of structures associated with flight control. We suspect that many of these sensory systems serve multiple roles. Some, such as antennae, may simultaneously provide information about gyroscopic sensing, gravity sensing and wind sensing, in addition to their classic role as chemosensory structures. Other structures (such as wings) have strain sensors that may serve a gyroscopic function as well as the measurement of aerodynamic loading. Preliminary neurobiological studies of these mechanosensory systems all point to extraordinary rates of information flow and high levels of precision in motion sensing.

Acknowledgments

This work was supported by grants to TLD from DARPA, ONR and the Joan and Richard Komen Endowed Chair. Support to JLF is via an NIH Training Grant in Neurobiology.

References

Berry MJ, 2nd, Warland DK, Meister M (1997) The structure and precision of retinal spike trains. Proc Natl Acad Sci U S A 94: 5411–5416

Borst A, Theunissen FE (1999) Information theory and neural coding. Nat Neurosci 2: 947–957

Burkhardt D, Schneider G (1957) Die Antennen von *Calliphora* als Anzeiger der Fluggeschwindigkeit. Z Naturf 12 b: 139–143

Chan WP, Prete F, Dickinson MH (1998) Visual input to the efferent control system of a fly's "gyroscope". Science 280: 289–292

Collett TS, Land MF (1975) Visual control of flight behaviour in the hoverfly *Syritta pipiens L.* J Comp Physiol A 99: 1–66

Combes SA, Daniel TL (2003) Flexural stiffness in insect wings. I. Scaling and the influence of wing venation. J Exp Biol 206: 2979–2987

Daniel TL, Combes SA (2002) Flexible wings and fins: bending by inertial or fluid-dynamic forces? Integr Comp Biol 42: 1044–1049

Daniel TL, Dieugonne A, Fox J, Myhrvold C, Sane S, Wark B (2008) Inertial guidance systems in insects: from the neurobiology to the structural mechanics of biological gyroscopes. J Inst Navigation 55: 235–240

Dickinson MH (1999) Haltere-mediated equilibrium reflexes of the fruit fly, *Drosophila melanogaster*. Philos Trans R Soc Lond B 354: 903–916

Dombrowski U (1991) Untersuchungen zur funktionellen Organisation des Flugsystems von *Manduca sexta* (L.) Universität Köln Dissertation

Etheredge JA, Perez SM, Taylor OR, Jander R (1999) Monarch butterflies (*Danaus plexippus* L.) use a magnetic compass for navigation. Proc Natl Acad Sci U S A 96: 13845–13846

Fayyazuddin A, Dickinson MH (1996) Haltere afferents provide direct, electrotonic input to a steering motor neuron in the blowfly, *Calliphora*. J Neurosci 16: 5225–5232

Fox JL, Daniel TL (2008) A neural basis for gyroscopic force measurement in the halteres of *Holorusia.* J Comp Physiol A 194: 887–897

Fratzl P, Barth FG (2009) Biomaterial systems for mechanosensing and actuation. Nature 462: 442–448

Frye MA (2001) Encoding properties of the wing hinge stretch receptor in the hawkmoth *Manduca sexta*. J Exp Biol 204: 3693–3702

Frye MA (2007) Behavioral neurobiology: a vibrating gyroscope controls fly steering maneuvers. Curr Biol 17: R134–6

Gewecke M (1970) Antennae: another wind-sensitive receptor in locusts. Nature 225: 1263–1264

Gewecke M, Niehaus M (1981) Flight and flight control by the antennae in the Small Tortoiseshell (*Aglais urticae* L., Lepidoptera). J Comp Physiol A 145: 249–256

Hedrick TL, Daniel TL (2006) Flight control in the hawkmoth *Manduca sexta*: the inverse problem of hovering. J Exp Biol 209: 3114–3130

Heran H (1959) Wahrnehmung und Regelung der Flugeigengeschwindigkeit bei *Apis mellifica L.* Z Vergl Physiol 42: 103–163

Hinterwirth AH, Daniel TL (2010) Antennae in the hawkmoth Manduca sexta (Lepidoptera, Sphingidae) mediate abdominal flexion in response to mechanical stimuli. J Comp Physiol A 196(12): 887–897

Huston SJ, Krapp HG (2009) Nonlinear integration of visual and haltere inputs in fly neck motor neurons. J Neurosci 29: 13097–13105

Kamikouchi A, Inagaki HK, Effertz T, Hendrich O, Fiala A, Göpfert MC, Ito K (2009) The neural basis of *Drosophila* gravity-sensing and hearing. Nature 458: 165–171

Keil TA (1997) Functional morphology of insect mechanoreceptors. Microscopy Research and Technique 39: 506–531

Kloppenburg P, Camazine SM, Xue JS, Randolph P, Hildebrand J (1997) Organization of the antennal motor system in the sphinx moth *Manduca sexta*. Cell Tissue Res 287: 425–433.

Köppl C (1997) Phase locking to high frequencies in the auditory nerve and cochlear nucleus magnocellularis of the barn owl, *Tyto alba*. J Neurosci 17: 3312–3321

Land MF, Collett TS (1974) Chasing behaviour of houseflies (*Fannia canicularis*). J Comp Physiol 89: 331–357

Lehmann F-O (2009) The limits of turning control in flying insects. In: Floreano D, Zufferey J-C, Srinivasan MV, Ellington C (eds) Flying insects and robots. Springer, Berlin, pp 231–246

Mountcastle A, Daniel T (2009) Aerodynamic and functional consequences of wing compliance. Experiments in Fluids 46: 873–882

Nalbach G (1993) The halteres of the blowfly *Calliphora* I. Kinematics and dynamics. J Comp Physiol A 173: 293–300

Niehaus M (1981) Flight and flight control by the antennae in the Small Tortoiseshell (*Aglais urticae* L., Lepidoptera). J Comp Physiol A 145: 257–264

Pix W, Nalbach G, Zeil J (1993) Strepsipteran forewings are haltere-like organs of equilibrium. Naturwissenschaften 80: 371–374

Pringle JWS (1948) The gyroscopic mechanism of the halteres of Diptera. Phil Trans R Soc Lond B BiolSci 233: 347–384

Reiser MB, Dickinson MH (2008) A modular display system for insect behavioral neuroscience. J Neurosci Methods 167: 127–139

Rieke F, Warland D, Bialek W, de Ruyter van Steveninck R (1997) Spikes: exploring the neural code. MIT Press, Cambridge, MA; London

Sandeman DC (1980) Angular acceleration, compensatory head movements and the halteres of flies (*Lucilia serricata*). J Comp Physiol A 136: 361–367

Sane SP, Dieudonne A, Willis MA, Daniel TL (2007) Antennal mechanosensors mediate flight control in moths. Science 315: 863–866

Sane SP, McHenry MJ (2009) The biomechanics of sensory organs. Integr Comp Biol 1–16.doi: 10 1093/icb/icp112.

Sprayberry JD (2009) Responses of descending visually-sensitive neurons in the hawkmoth, *Manduca sexta*, to three-dimensional flower-like stimuli. J Insect Sci 9: 7

Srinivasan MV (2006) Small brains, smart computations: Vision and navigation in honeybees, and applications to robotics. Brain-Inspired IT II. International Congress Series, vol 1291, pp 30–37

Srinivasan MV, Zhang S, Lehrer M, Collett T (1996) Honeybee navigation en route to the goal: vis-

ual flight control and odometry. J Exp Biol 199: 237–244

Theobald JC, Warrant EJ, O'Carroll DC (2009) Wide-field motion tuning in nocturnal hawkmoths. Proc Biol Sci doi: 10 1098/rspb.2 009 1677

Theunissen F, Roddey JC, Stufflebeam S, Clague H, Miller JP (1996) Information theoretic analysis of dynamical encoding by four identified primary sensory interneurons in the cricket cercal system. J Neurophysiol 75: 1345–1364

Thompson RA, Wehling MF, Evers JH, Dixon WE (2009) Body rate decoupling using haltere mid-stroke measurements for inertial flight stabilization in Diptera. J Comp Physiol A 195: 99–112

Warzecha A, Egelhaaf M (2000) Response latency of a motion-sensitive neuron in the fly visual system: dependence on stimulus parameters and physiological conditions. Vision Res 40: 2973–2983

Infrared and electro-reception V

Designing a fluidic infrared detector based on the photomechanic infrared sensilla in pyrophilous beetles

20

Helmut Schmitz, Herbert Bousack

Contents

Abstract

Beetles of the genus *Melanophila* approach forest fires from large distances away. For the detection of fires and of hot surfaces beetles have evolved infrared (IR) receptors, which have developed from common hair mechanoreceptors and are therefore termed photomechanic. Compared to a hair mechanoreceptor, a photomechanic IR sensillum shows the following two special features: (i) The formation of a complex cuticular sphere instead of the hair shaft; the sphere consists of an outer exocuticular shell as well as of a cavernous microfluidic core. (ii) The enclosure of the dendritic tip of the mechanosensitive neuron inside the core in a liquid-filled chamber. Hence we propose that the photomechanic IR sensillum represents a microfluidic converter of infrared radiation to an increase of the pressure inside the sphere, which is in turn measured by a mechanosensitive neuron.

A simple model of this biological IR sensor is the Golay sensor filled with a liquid instead of gas. Here absorption of IR radiation results in a pressure increase of the liquid and the deflection of a thin sensor membrane. For the evaluation of this model analytical formulas are presented for the calculation of the pressure increase in the cavity, the deformation of the membrane and the time constant of an artificial leak to compensate for ambient temperature changes. Some organic liquids with high thermal expansion coefficients may improve the deflection of the membrane as compared to water.

Helmut Schmitz
University of Bonn, Institute of Zoology
Poppelsdorfer Schloß, 53115 Bonn, Germany
e-mail: h.schmitz@uni-bonn.de

1. Introduction

"Fire-loving" (pyrophilous) beetles depend on forest fires for their reproduction. They approach ongoing fires and invade the burnt area immediately. For long-range navigation towards a fire as well as for short-range orientation on a freshly burnt area these beetles have special sensors for smoke and infrared (IR) radiation. Whereas the olfactory receptors for smoke are located on the antennae, the IR receptors are housed in extraantennal sensory organs found on the thorax or on the abdomen. In the best-studied pyrophilous beetle, *Melanophila acuminata*, infrared receptors and their associated sensory neurons are derived from mechanoreceptors (Schmitz et al. 2007). Unlike other mechanosensory neurons, IR sensitive neurons immediately send their information to be processed centrally by the brain rather than locally in their respective ganglia of origin (Gronenberg and Schmitz 1999). It is suggested that smoke derived odours and IR information converge on descending brain neurons which, in turn, control and direct flight toward the forest fire.

Two genera of jewel beetles (family Buprestidae) can be classified as pyrophilous: About a dozen species of the genus *Melanophila* that are distributed all over the world except for Australia and the "fire-beetle" *Merimna atrata* that is endemic to Australia (Poulton 1915; Schmitz and Schmitz 2002). On the freshly burnt area, the males of both genera often stay on the stems of trees close to burning or glowing wood or hot ashes. As soon as they become aware of a conspecific female, they try to copulate vigorously. After mating, the females deposit the eggs under the bark of burnt trees. The main reason for the pyrophilous behavior is that the wood-boring larvae of *Melanophila* and *Merimna* can only develop in the wood of burnt trees (Apel, 1989; Schmitz and Schmitz 2002). As a morphological speciality both pyrophilous buprestid genera are equipped with antennal smoke receptors and thoracic or abdominal IR organs (Schmitz and Bleckmann 1997; Schütz et al. 1999; Schmitz et al. 2001; Schmitz and Trenner 2003).

Another pyrophilous beetle can be found in Australia: the "little ash beetle" *Acanthocnemus nigricans* (family Acanthocnemidae). This inconspicuous beetle is only 4 mm long and strongly attracted by hot ashes. However, its biology is nearly unknown. Obviously, *Acanthocnemus* also depends on fires for its reproduction and is equipped with a pair of sophisticated prothoracic IR receptors (Schmitz et al. 2002; Kreiss et al., 2005). Up to now, nothing is known about possible smoke detectors in *Acanthocnemus*.

Fire detection is the compelling precondition for the survival of all pyrophilous beetle species mentioned above. However, the outbreak of a forest fire is highly unpredictable. In the northern boreal forests, where most *Melanophila* species can be found, a fire statistically occurs only every 50–200 years. Because finding a fire is crucial for the survival of all *Melanophila* species, *Melanophila* beetles need to be able to detect fires from distances as large as possible. Therefore, it has to be postulated that the sensory organs which are used for fire detection must have been subjected to a particularly strong evolutionary pressure with regard to sensitivity. Furthermore, when flying over a burnt area in search of a landing ground, pyrophilous insects have to avoid "hot spots" with dangerous surface temperatures above about 60 °Celsius. Finally, the fire-relevant information provided by the peripheral sensors has to be processed by a highly efficient central neuronal network guiding the insect straight to a fire.

2. The photomechanic IR receptors in pyrophilous *Melanophila* beetles

2.1 Structure

Buprestid beetles of the genus *Melanophila* inhabit the palaearctic as well as the nearctic regions of the world. Both sexes approach forest fires because their life cycle depends on burnt wood as larval food (Champion 1909; Linsley 1943; Evans 1962; Apel 1989; Gronenberg 1996). Initially, the freshly burnt area serves as a meeting place where the males search for females. For the detection of fires, *Melanophila* beetles are equipped with special antennal smoke detectors (Schütz et al., 1999) and thoracic IR receptors (Schmitz and Bleckmann 1997). The single IR receptors (called sensilla in insects) are situated in two pit organs located on the third thoracic segment. Each IR organ houses about 70 IR sensilla which are closely packed together at the bottom of the pit (Vondran et al. 1995) (see Fig. 1).

From the outside, a single sensillum can be recognized by a hemispherical dome with a diameter of about 12–15 µm (Fig. 2). The dome is built by a thin cuticle which represents the outer boundary of a spherical internal cavity. The cavity is almost completely filled by a tiny cuticular sphere with a diameter of about 10 µm. Based on transmission electron microscopical observations (Vondran et al. 1995) the sphere consists of three different zones: an outer lamellated mantle, an intermediate layer of unstructured cuticle revealing many irregularly arranged microcavities, and an innermost zone where the cuticle appears uniform except for some spots of higher electron density. The sphere is connected to the vertex of the outer cuticular dome by a small cuticular stalk and the narrow gap surrounding the sphere is filled by leaf like extensions of at least two non-neural enveloping cells. From below, the sphere is innervated by a single sensory cell. All morphological as well as all physiological data available has so far demonstrated that this cell is a ciliary mechanoreceptor (Vondran et al., 1995; Schmitz et al., 2007).

2.2 Function

Currently, two models of sensillum function can be found in the literature, but which are inconsistent with each other. (i) In the first one a photomechanic principle was established (Schmitz et al., 1997; Schmitz and Bleckmann 1998) proposing that the biomolecules (i. e. proteins and chitin) of the cuticular sphere strongly absorb mid-IR radia-

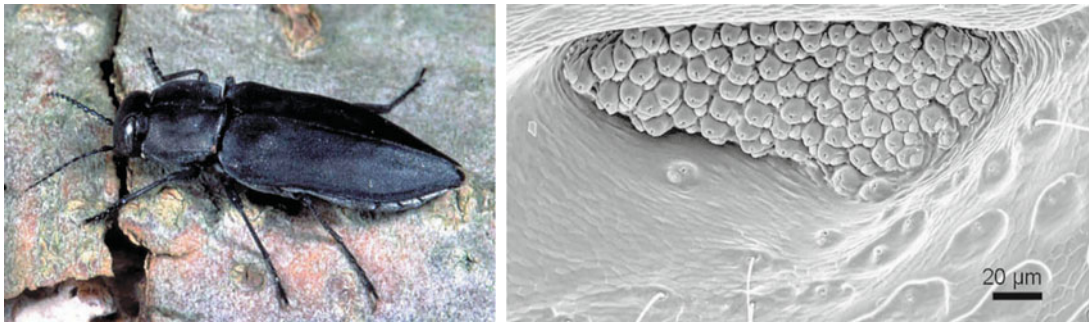


Fig. 1 Left: The "firebeetle" *Melanophila acuminata* is about 1 cm long. On either side of the body one IR organ is found directly behind the bases of the middle legs. **Right:** One IR pit organ containing about 70 hemispherical IR sensilla

tion. In a way not understood in detail, the resulting thermal expansion of the cuticular sphere is thought to be measured by the mechanosensitive receptor cell. (ii) The second model assumes the presence of an air-filled cavity inside the sphere (Evans 2005) and suggests that IR radiation enters this cavity by a small apical waveguide (diameter about 1.5 µm). Due to the absorption of IR photons at the inner cuticular walls of the cavity the enclosed air is said to be heated up and to expand. In a way not further specified the resulting increase in gas pressure is believed to stimulate the mechanoreceptor.

According to our recent findings the thoracic infrared (IR) sensilla of the pyrophilous jewel beetle *Melanophila acuminata* most likely evolved from hair mechanoreceptors (sensilla trichodea) (Schmitz et al. 2007). Hair mechanoreceptors, which can be found in any insect, have cuticular hair shafts of different length. In insects the strain–sensitive dendrite of a single mechanosensitive sensory cell is attached to the base of the hair shaft. As a result, even a slight deflection of the hair shaft stimulates the tip of the dendrite. To further elucidate the sensory transduction mechanism, the morphology of IR sensilla and of neighbouring hair mechanoreceptors was investigated by using electron microscopical techniques (SEM, TEM) in combination with focused ion beam milling (FIB). It was assumed that any deviation from the *bauplan* of a hair mechanoreceptor is of particular relevance for the transformation of IR radiation to a mechanical stimulus. This makes even more sense in the light of the homology of sensor structures presumed to be relevant for stimulus uptake and transduction. Compared to a hair mechanoreceptor, an IR sensillum shows the following special features (see Fig. 2). First, a complex cuticular sphere is formed instead of the hair shaft; the sphere consists of an outer exocuticular shell as well as of an inner porous mesocuticular part. Second, the dendritic tip of the mechanosensitive neuron inside the sphere is enclosed in a fluid-filled larger basal chamber, which is connected to a system of microcavities and nanocanals in the mesocuticular part (Schmitz et al. 2007). Most probably, IR radiation absorbed by the proteins, the chitin fibres and the fluid of the sensillum heats up the sphere causing immediate thermal expansion especially of the fluid inside the spongy mesocuticular layer. According to the morphological results of Schmitz et al. (2007) the only compliant structure in the sphere is the membrane of the tip of the mechanosensitive dendrite. It is reasonable to suppose that the deflection of the membrane by a few nanometers opens stretch-activated ion channels. Hence we propose that an IR sensillum represents a microfluidic converter of infrared radiation into an increase of the pressure inside the sphere which is then measured by the mechanosensitive neuron. The proposed principle of transforming IR radiation into a bioelectrical signal has been termed photomechanic. The existence of a central air-filled cavity in the sphere as described by Evans (2005) could not be verified. We therefore exclude

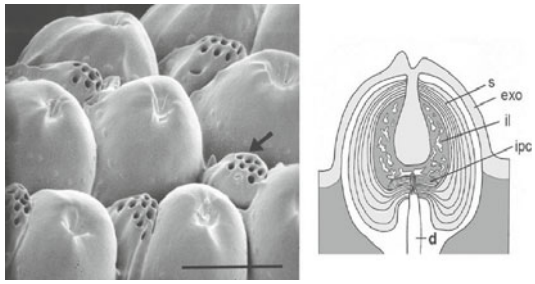


Fig. 2 Left: The hemispherical IR sensilla located at the bottom of the pit organ. Each of the small dome-shaped receptors (diameter 12–15 µm) is accompanied by a wax gland (arrow). Bar 10 µm. **Right:** Semischematic drawing of a 2-dimensional section through the middle of an IR sensillum. exo: stiff exocuticle; s: outer exocuticular shell of internal sphere. Lamellation is caused by concentric layers of reinforcing chitin fibres

the possibility that the *Melanophila* sensillums relies on the expansion of gas and can thus be classified as a biological Golay cell.

3. Golay sensor

A simple model of this biological IR-Sensor is the pneumatic Golay sensor (Golay 1947). It consists of an internal gas-filled cavity, which is closed on one side by a window and on the other side by a thin membrane. IR-radiation enters through the window and heats up the gas by absorption. The deflection of the membrane caused by the expanding gas can be read by an optical system (Golay 1947), a capacitive detector (Chevrier et al. 1995), or a tunneling displacement transducer (Kenny 1996). To enhance the IR-absorption in the gas the cavity is equipped with an additional absorber. Reflecting walls of the cavity are another way to enhance the absorption. The detectivity D* (Jones 1960) is a figure of merit to characterize the performance of IR detectors. D* is defined as the radiation power, normalized to the unit area and unit bandwidth, required to get a signal from the IR detector equal to the detector noise. A conventional Golay sensor D* is about $4 \cdot 10^9\,\mathrm{cmW^{-1}\,Hz^{1/2}}$ (Golay 1949). Such sensors micromachined from silicon have been designed by several authors (Chevrier et al. 1995; Kenny 1996; Ajakaiye et al. 2007).

4. Sensor components

4.1 Fluid

The most effective sensor is one which shows the least heat loss of the cavity. M. J. Golay recommended using the least conductive gas available, xenon (Golay 1949). The use of water instead of gas has some consequences. Compared to xenon or carbon dioxide the thermal conductivity of water is larger by a factor of 100 or 40, respectively. This means that the heat loss of a water-filled cavity is noticeably higher. The time constant τ is a measure of the response time of a sensor. A sensor with a low value of τ is reacting quickly. τ is defined as

$$\tau = \Theta \cdot \frac{\rho \cdot c_P}{\lambda} \qquad (1)$$

where Θ is a factor depending on geometry and boundary conditions, ρ, density, c_p, heat capacity, and λ, heat conductivity of the fluid inside the cavity.

Due to its high density and heat capacity compared to gas ($\tau_{water}/\tau_{CO2} \cong 45$), water as a cavity fluid will reduce the speed of the sensor's response. An advantage of water is that its IR absorption is very high (for example about 5 orders of magnitude higher than that of carbon dioxide) and therefore no additional IR-absorber will be necessary. The detectivity D* of an IR-sensor is one possible figure of merit to rate these influencing factors. Bousack and Offenhäuser (2007) compared the effects of various fluids and the geometry of the cavity on the detectivity of the sensor.

4.2 Cavity insulation

A way to reduce the heat loss of the cavity and to thereby achieve a higher fluid temperature is to improve the insulation of the

cavity. The exocuticular shell and the inner spongy layer of mesocuticle of the sensilla act as an insulation of the inner pressure chamber because the thermal conductivity of cuticle is low. The thermal effusivity *b* is a measure to quantitatively describe and compare the heat insulation properties of different materials.

$$b = \sqrt{\lambda \cdot \rho \cdot c_P} \qquad (2)$$

For a sensor produced with microsystems technology, silicon is commonly used. Silicon however has a high thermal effusivity of 16 300 $Ws^{1/2}/m^2$ K compared to the thermal effusivity of cuticle, 400–650 $Ws^{1/2}/m^2K$. That means that silicon is less suitable for a good thermal insulation of the cavity. Plastics have a thermal effusivity of about 300–900 $Ws^{1/2}/m^2K$, which is comparable to that of cuticle. Drawbacks and opportunities using plastics for the cavity must be analyzed. In the best-case scenario, the cavity is fully thermally insulated. In this case, the change in temperature inside the cavity can be calculated from simple energy balance considerations:

$$T(t) = T_0 + \frac{I_0 \cdot A}{m \cdot c_P} \cdot t \qquad (3)$$

where T, T_0 are the mean and initial temperature inside the cavity, I_0 is the IR radiation density, A the cross section surface of the cavity, m the mass of fluid inside the cavity and t the time.

Because the temperature progress is linear, the IR radiation density I_0 can be calculated immediately without time restrictions due to a large time constant. Perhaps this explains why *Melanophila acuminata* uses a slow sensor with a high thermal mass but with fast response time of some milliseconds. In addition the time constant of the *Melanophila acuminata* sensor is reduced because of its small dimensions.

4.3 Cavity shape and size

Regarding an effective energy balance of the cavity, the surface of the IR-window should be large (gathering more IR) and the volume of the cavity should be small (reducing energy losses across the surface and causing a larger temperature increase due to a smaller thermal mass). For this reason a small ball-shaped cavity would be a good solution. It is not surprising that nature chose such a shape for the sensillum. A technical solution manufactured using microsystems technology, however, needs to have a cylindrical shape. For a cylinder the energy balance is less favorable than for a sphere, especially when its length (depth of the cavity) is larger than its diameter. An additional limitation is the diameter of the cavity, if its bottom forms a thin membrane, whose deflection is used as a read-out for the pressure signal.

4.4 Sensor read-out

The sensillum senses the deformation (strain) of the dendritic membrane. For a technical read-out the deflection of a thin membrane at the bottom of the cavity can be used. The deflection z of a circular membrane due to a pressure difference across its two surfaces can be calculated as a function of the radial distance applying shell theory (Szabo 2001).

$$z(r) = \frac{\Delta P}{64 \cdot D} \cdot \left(R^2 - r^2\right)^2 \qquad D = \frac{E \cdot t_P^3}{12 \cdot (1 - \nu^2)} \qquad (4)$$

where ΔP is the pressure difference, R the radius of the membrane, D its flexural stiffness, E Young's modulus of elasticity, t_P the thickness of the membrane, and *v* Poisson's ratio. This formula is only applicable for $z_{max} < 1 \cdot t_P$. When the diameter of the cavity is reduced, the membrane must obviously become thinner to be sufficiently deflected. This ultimately puts a limit on the minimum diameter of the cavity. The actual deflec-

tion of the membrane can be measured by optical methods, e. g. interferometry, piezoresisitive films on the membrane or by a capacitor, using the membrane as one of its two plates. For a capacitor the change of the capacitance due to a pressure increase in the cavity can be calculated (Bousack and Offenhäusser 2007), assuming a constant charge distribution on the capacitor plates.

$$C_K = \varepsilon_0 \cdot \varepsilon_r \cdot 2 \cdot \pi \int_0^R \frac{r}{d_0 - z_{(r)}} \cdot dr$$

$$C_K = \frac{\varepsilon_0 \cdot \varepsilon_r \cdot \pi}{\sqrt{\frac{\Delta P \cdot d_0}{64 \cdot D}}} \cdot \operatorname{arctan\,h} \sqrt{\frac{\Delta P \cdot R^4}{64 \cdot D \cdot d_0}} \quad (5)$$

where C_K is the capacitance, ε_0 the permittivity of free space, ε_r the relative permittivity, and d_0 the distance between the capacitor plates.

4.5 Compensation leak

In the *Melanophila* IR receptor the inner sphere is enclosed by a thin layer of fluid. The nanocanals in the shell of the sphere allow the exchange of fluid in and out of the microfluidic compartment in the sphere (Schmitz et al. 2007). Thus, any internal pressure change which may be caused by slowly changing ambient temperature can be compensated for. Golay sensors also use such compensation leaks for compensating changes of ambient temperatures (Chevrier et al 1995). The geometric design of such a compensation leak for a cavity was analyzed in Bousack et al. (2008).

5. Sensor model

A technical version of the sensor is shown in Fig. 3. Similar to the sensillum the sensor contains an internal water-filled cavity, which is etched into a silicon wafer. The cavity is closed on one side by a window made from Suprasil (Heraeus 2009) and on the other side by a thin silicon membrane. The absorbed IR-radiation produces a change in pressure and the deflection of the membrane caused by this pressure increase can be read out by a capacitive detector.

For the calculation of the temperature distribution inside the cavity a simple assumption is a uniformly distributed power density using the Lambert-Beer law of absorption.

$$q_e''' = \frac{I_0 \cdot (1 - e^{-\mu \cdot H_C})}{H_C} \quad (6)$$

where I_0 is the IR radiation density at z = 0, μ the absorption coefficient, and H_C the height of the cylindrical cavity. For the uniformly distributed power density the temperature distribution, the deflection of the membrane and the capacitor signal can be calculated analytically (Bousack and Offenhäusser 2007) assuming an isochoric change of state. In reality the absorption takes place in a very thin zone where the absorption co-

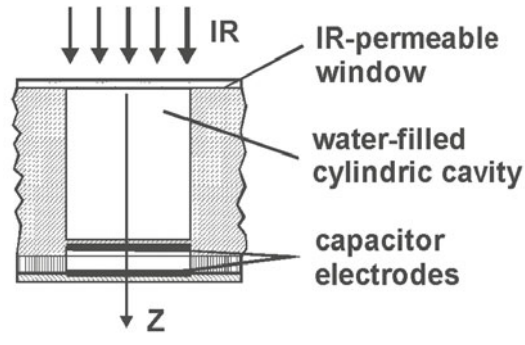


Fig. 3 Technical model of the sensor

efficient is very high as in case of water. In this case the power density is a function of the coordinate z.

$$q(z)''' = I_0 \cdot \mu \cdot e^{-\mu \cdot z} \qquad (7)$$

It is obvious that a significant absorption and heat production inside the cavity occurs only in a very thin zone of the cavity directly behind the window.

To assess the influence of power density with respect to equations (6) and (7) the temporal temperature distribution was calculated for the model in Fig. 3 using the software COMSOL Multiphysics® for the 3-dimensional case. In this example a window made from Suprasil (window thickness 0.5 mm), a cavity depth and diameter of 0.5 mm in silicon, and an IR density I_0= 10 W/m^2 at the outer window surface are assumed, using a mean absorption coefficient μ in a wavelength window 2–4 μm of 50 1/m for Suprasil and 2.2 · 10^5 1/m for water (Wieliczka et al. 1989). The absorption coefficient μ is a measure of the amount of absorbed IR energy in a medium per unit-length in a wavelength window, see equations (6) and (7). Evidently, a significant amount of the heat produced in the absorption zone of the cavity is conducted into the window due to the approximately two-times larger heat conductivity of the window material compared to that of water. The temperature profiles in Fig. 4 indicate that in the case of a thin absorption zone only 50–60 % of the energy compared to the uniformly distributed power density is stored in the cavity and that for times short after the start of the irradiation the maximum temperature is higher.

A reduction of the cavity depth e. g. to 0.2 mm may seem to be a reasonable proposal. However, manufacturing problems using thin wafers, the water filling procedure etc. have to be investigated first.

For the calculation of the pressure in the cavity based on the temperature profile in the absorption zone the equation of state must be solved. Because the pressure depends upon the two independent variables temperature and volume, the total derivative gives

$$\Delta P = \beta \cdot \Delta T_{mean} - \frac{1}{\kappa \cdot V} \cdot \Delta V \qquad (8)$$

Where Δ*P* is the pressure difference, $\beta = (\partial P/\partial T)_V$ the isochoric tension coefficient, $\kappa = -(1/V) \cdot (\partial V/\partial P)_T$ the isothermal compress-

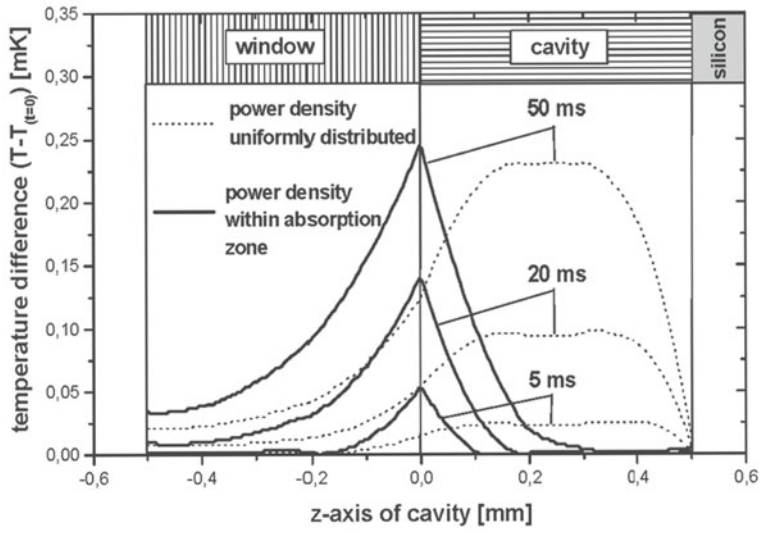


Fig. 4 Temperature distribution after 5, 20 and 50 ms irradiation (radiation density of 10 W/m^2) at the outer surface of the window using a uniformly distributed power density, see equation (6), and a power density due to an absorption zone, see equation (7)

ibility coefficient, and V the volume of the cavity. For water (25 °C, 1 bar) $\beta = 5.68$ bar/K and $\kappa = 45 \cdot 10^{-6}$ 1/bar (Weast 1981).

One can assume that the increase in volume due to the deflection of the membrane is negligible (isochoric change of state), that means $(\Delta V/V)/\kappa << \beta \cdot \Delta T$ in equation (8). With the mean temperature calculated from the temperature profile (uniformly distributed power density in the cavity in Fig. 4) at t = 50 ms a pressure increase of 110 N/m^2 results. For a more precise calculation of the pressure increase the temperature profile of the thin absorption zone, see Fig. 4, and the influence of the second term in equation (8) will be considered. The increase in the volume ΔV of the cavity is due to the tiny deflection of the membrane. Using equation (4) the volume ΔV can be calculated in relation to the pressure increase ΔP.

$$\Delta V = \frac{\pi \cdot R^6 \cdot \Delta P \cdot \left(1 - \upsilon^2\right)}{16 \cdot E \cdot t_P^3} \qquad (9)$$

Combining equations (8) and (9) yields a relationship for ΔP which considers the influence of temperature and volume.

$$\Delta P = \frac{\beta \cdot \frac{\int_0^{H_C} \Delta T_{(z)} \cdot dz}{H_C}}{1 + \Omega} \qquad (10)$$

$$\Omega = \frac{R^4 \cdot (1 - \nu^2)}{16 \cdot E \cdot t_P^3 \cdot H_C \cdot \kappa}$$

For water as a fluid with a high IR absorption coefficient a formula for the temperature profile $\Delta T_{(z)}$ could be derived (Soltner et al 2009), which allows a numerical solution of equation (10). Because the value of Ω in equation (10) is larger than unity, the pressure increase ΔP is reduced considerably in comparison to the isochoric change of state (Ω << 1) mentioned above. An isobaric change of state means ΔP = 0 because of $\Omega \rightarrow \infty$ (very soft membrane), in this case the deflection of the "membrane" is only due to the thermal expansion of the water. In reality neither the isochoric nor the isobaric change of state is feasible. To achieve high pressure in the cavity or a large membrane deflection a complex optimisation is necessary.

Conclusions

Like the photomechanic insect IR receptors, our technical model features very small displacements of the sensor membrane in the range of 1 nm and less. The insects have the advantage of possessing a very sensitive mechanoreceptor, the mechanosensitive neuron, which can detect minute deflections (strain) of its membrane. It seems that this sensitivity of the beetle is difficult to achieve in an inexpensive and robust technical sensor.

For our miniaturized technical sensor, an appropriate technical read-out mechanism with a high resolution of 1 nm may be necessary. So far a sensor with a capacitor as read-out is near to completion in the *Center of Advanced European Studies and Research* (caesar), Bonn. Several read-out methods will be evaluated, but it has to be stressed that this search must yield a rugged and cost-effective design in order to be able to compete with existing IR sensors.

An obvious method for enhancing the displacement of the membrane is to choose liquids with optimal thermal properties, e. g. high thermal expansion coefficient and low heat capacity. Here several hydrocarbon liquids commonly used in expansion thermometers permit an improvement. Methanol is a good candidate compared to n-pentane and toluene due to better handling during the filling process of the cavity. Additionally care has to be taken in selecting appropriate wall materials with a low thermal conduction necessary to reduce the heat loss of the cavity (at the cost of the sensor's temporal resolution). Until now silicon is used because of its well known manufacturing technology for micro systems. As an alternative micro injection molding of plastic materials with low heat conductivity is under investigation.

Acknowledgements

The authors acknowledge useful discussions with H. Soltner, Forschungszentrum Jülich. H. Bleckmann, University of Bonn, provided lab space. Supported by a grant from the German Ministry of Education and Research (BMBF).

References

Ajakaiye O, Grade J, Choongsoo S, Kenny T (2007) Wafer-scale fabrication of infrared detectors based on tunneling displacement transducers. Sensor Actuator A 134: 575–581

Apel K-H (1989) Zur Verbreitung von *Melanophila acuminata* DEG. (Col., Buprestidae). Entomol Nach Ber 33: 278–280

Bousack H, Offenhäusser A (2007) Systemanalyse eines fluidischen Infrarot-Detektors mit kapazitiver Auslese als einfaches Modell für die infrarotempfindlichen Rezeptoren des Schwarzen Kiefernprachtkäfers. Proc MikroSystemTechnik Kongr Dresden: 991–994

Bousack H, Schmitz H, Offenhäusser A (2008) Design of a fluidic infrared detector based on the infrared sensilla in the beetle *Melanophila acuminata*. Actuator 2008, 11th Int Conference on New Actuators, Bremen: 786–789

Champion GC (1909) A buprestid and other Coleoptera on pine injured by "heath fires" in N. W. Surrey. Entomol Mon Mag 45: 247–250

Chevrier J, Baert K, Slater T (1995) An infrared pneumatic detector made by micromachining technology. J Micromech Microeng 5: 193–195

Evans WG (1962) Notes on the biology and dispersal of *Melanophila* (Coleoptera: Buprestidae). Pan-Pacific Entomol 38: 59–62

Evans WG (2005) Infrared radiation sensors of *Melanophila acuminata* (Coleoptera: Buprestidae): A thermopneumatic model. Ann Entomol Soc Am 98(5): 738–746

Golay M (1947) A pneumatic infra-red detector. Rev Sci Instrum 18(5): 357–362

Golay M (1949) The theoretical and practical sensitivity of the pneumatic infra-red detector. Rev Sci Instrum 20(11): 816–820

Gronenberg W (1996) The trap-jaw mechanism in the dacetine ants *Daceton armigerum* and *Strumigenys sp*. J Exp Biol 199: 2021–2033

Gronenberg W, Schmitz H (1999) Afferent projections of infrared sensilla in the beetle *Melanophila acuminata* (Coleoptera: Buprestidae). Cell Tissue Res 297: 311–318

Jones R C (1960) Proposal of the specific detectivity D* for detectors limited by radiation noise. J Op Soc Am 50 (11): 1058–1059

Kenny T (1996) Micromachined infrared sensors using tunneling displacement transducers. Rev Sci Instrum 67(1): 112- 128

Kreiss E-J, Schmitz A, Schmitz H (2005) Morphology of the prothoracic discs and associated sensilla of *Acanthocnemus nigricans* (Coleoptera, Buprestidae). Arthropod Struct Dev 34: 419–428

Linsley EG (1943) Attraction of *Melanophila* beetles by fire and smoke. J Econ Entomol 36: 341–342

Poulton EB (1915) The habits of the Australian buprestid "fire-beetle" *Merimna atrata*, Lap. et Gory. Trans Entomol Soc London pt. 1, proc.: iii–iv

Schmitz A, Sehrbrock A, Schmitz H (2007) The analysis of the mechanosensory origin of the infrared sensilla in *Melanophila acuminata* (Coleoptera; Buprestidae) adduces new insight into the transduction mechanism. Arthropod Struct Dev 36: 291–303

Schmitz H, Bleckmann H (1997) Fine structure and physiology of the infrared receptor of beetles belonging to the genus *Melanophila* (Coleoptera: Buprestidae). Int J Insect Morphol & Embryol 26: 205–215

Schmitz H, Bleckmann H (1998) The photomechanic infrared receptor for the detection of forest fires in the buprestid beetle *Melanophila acuminata*. J Comp Physiol A 182: 647–657

Schmitz H, Mürtz M, Bleckmann H (1997) Infrared detection in a beetle. Nature 386: 773–774

Schmitz H, Schmitz A (2002) Australian fire-beetles. Landscape Spring 2002: 36–41

Schmitz H, Schmitz A, Bleckmann H (2001) Morphology of a thermosensitive multipolar neuron in the infrared organ of *Merimna atrata* (Coleoptera, Buprestidae). Arthropod Struct Dev 30(2): 99–111

Schmitz H, Schmitz H, Trenner S, Bleckmann H (2002) A new type of insect infrared organ of low thermal mass. Naturwissenschaften 89: 226–229

Schmitz H, Trenner S (2003) Electrophysiological characterization of the multipolar thermoreceptors in the "fire-beetle" *Merimna atrata* and comparison with the infared sensilla of *Melanophila acuminata* (both Coleoptera, Buprestidae). J Comp Physiol A 189: 715–722

Schütz S, Weissbecker B, Hummel HE, Apel K-H, Schmitz H, Bleckmann H (1999) Insect antennae as a smoke detector. Nature 398: 298–299

Soltner H, Bousack H, Schmitz H (2009) Zur transienten Temperaturentwicklung in einem fluidischen Infrarot-Detektor auf der Basis der infrarotempfindlichen Sensillen des Schwarzen Kiefernprachtkäfers *Melanophila acuminata*. Proc MikroSystemTechnik Kongress, Berlin

Szabo I (2001) Höhere technische Mechanik: nach Vorlesungen. Springer-Verlag, Berlin

Vondran T, Apel K-H, Schmitz H (1995) The infrared receptor of *Melanophila acuminata* DE GEER (Coleoptera: Buprestidae): ultrastructural study of a unique insect thermoreceptor and its possible descent from a hair mechanoreceptor. Tissue Cell 27(6): 645–658

Weast RC (1981) Handbook of chemistry and physics. CRC-Press, Boca Raton

Wieliczka D, Wenig S, Querry M (1989) Wedge-shaped cell for highly absorbent liquids: infrared optical constants of water. Appl Optics 28(9): 1714–1719

Remote electrical sensing: detection and analysis of objects by weakly electric fishes

21

Gerhard von der Emde

Contents

Abstract

With a special electric organ, weakly electric fish emit electric signals, each of which builds up an electric field around them in the water. To sense locally occurring voltages caused by their own signals, the fish use thousands of electroreceptor organs distributed over their skin. Nearby objects are detected because they distort the self-produced electric field and thus project electric images onto the array of electroreceptors. By analysing these electric images, the fish can detect and localize objects in the water and recognize their electrical and spatial properties. Inspired by these remarkable capabilities of weakly electric fish, we designed technical sensor systems that can solve similar sensing problems. We applied the principles of active electrolocation by building devices that produce electrical current pulses and simultaneously sense local current densities. Depending on the specific task, sensors can be designed which (i) detect an object, (ii) localize it in space, (iii) determine its distance and (iv) measure its electrical properties. Our biomimetic sensor systems proved to be insensitive to environmental disturbances such as heat, pressure or turbidity. They can be used in a wide range of applications, such as material identification, remote distance measurements and medical diagnostics.

Gerhard von der Emde
University of Bonn, Institute of Zoology
Poppelsdorfer Schloß, 53115 Bonn, Germany
e-mail: vonderemde@uni-bonn.de

1. Electroreception in Nature

For orientation in the environment, animals employ different senses, depending on the physical conditions prevailing in a particular habitat and on the evolutionary history of the particular species. Most animals, including humans, mainly use vision to detect and localize objects and other environmental parameters. In most species, eyes are adapt-

ed to the light conditions existent during the active hours of the particular animal species.

If no or very little light is available when an animal is active, evolution has favoured the development of alternative senses. These can be hearing (owls, bats), olfaction (many fishes), touch (for example with whiskers), water wave detection (all fishes), or the electric sense (electroreceptive fishes). Often, a combination of several senses is employed that complement each other.

Electroreception, i. e., the detection of naturally occurring electric stimuli by animals with specialized electroreceptors in their skin, can only be found in animals that live in water and thus is always coupled to an aquatic medium. Many marine and freshwater fishes, with the important exception of most (but not all) teleosts, are electroreceptive (Bullock et al. 1982; Hopkins 2009). Detection of electrical signals coming from the environment and estimating the position of their source is called passive electrolocation. This is in contrast with *active* electrolocation in weakly electric fishes, which, in addition to electroreception, implies the active production and emission of electric signals with specialized electric organs.

Passive electrolocation requires the presence of environmental electrical fields in order to detect objects or other structures. This disadvantage is overcome in active electrolocation, where the animal itself is the source of the electrical energy used for object detection (Lissmann and Machin 1958). Active electrolocation is used by so-called weakly electric fishes, which consist of two distantly related clades of fishes that live in African and South American tropical freshwater rivers, creeks or lakes (von der Emde 1998). With a specialized electric organ in their tail, these fishes can emit electrical signals, or electric organ discharges (EODs), into their surroundings. The self-produced electrical signal builds up an electrical field around the animal, which is perceived by an array of cutaneous electroreceptor organs that are distributed over the body surface of the fish. Objects are detected because they interact with and modulate the EOD and thus cause local field distortions that are detected by the animal's electroreceptor organs (von der Emde et al. 2008a).

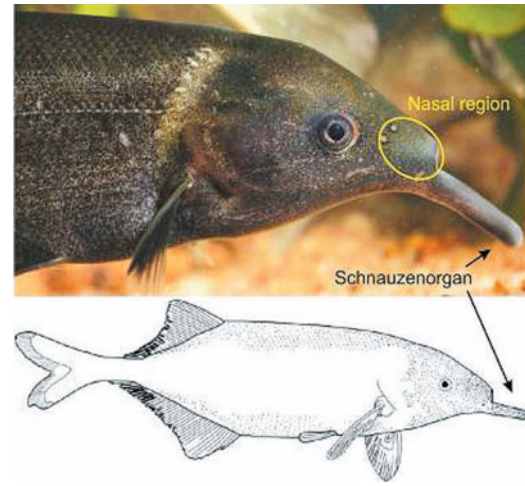


Fig. 1 *Gnathonemus petersii*, a member of the African weakly electric Mormyridae family. The electroreceptor organs used for electrolocation (Mormyromasts) are visible as small dots at the head and the back of the fish. The drawing at the bottom depicts the distribution of mormyromasts along the body. They occur at highest density in the nasal region and on the elongated and moveable chin appendix (Schnauzenorgan). The drawing of electroreceptor distributions was kindly provided by Dr. Michael Hollmann

In this chapter I shall elucidate the principal mechanisms of active electrolocation by using the African mormyrid fish *Gnathonemus petersii,* also known as the elephantnose fish, as an example (Fig. 1). This weakly electric fish is probably the best studied animal with respect to its active electrolocation system. *G. petersii* grows to a length of more than 30 cm and lives in freshwater rivers of Central and West Africa. It is mainly active at night when it orients itself in its habitat and finds its prey (small insect larvae living on the ground) primarily by employing

active electrolocation (von der Emde and Bleckmann 1998). A foraging *G. petersii* detects prey mainly with its movable chin appendix (called the *Schnauzenorgan*), which is covered with more than 500 epidermal electroreceptor organs (Hollmann et al. 2008).

2. Active electrolocation in weakly electric fish

During active electrolocation, modulation of the EOD is the carrier of environmental information for the animal. The modulations of the EOD are caused by objects close to the animals' skin differing in impedance from the surrounding water. Depending on the impedance difference between object and water, the current flowing locally through those electroreceptors apposing the object is either decreased or increased, resulting in a decrease or increase of the local EOD amplitude. Capacitive properties of nearby objects can additionally alter the waveform of the local EOD (von der Emde et al. 2008a). An "electric image" originating from the objects is thus projected onto the animal's electroreceptive skin, which typically has a centre-surround (or Mexican hat) spatial profile (Caputi et al. 1998; von der Emde et al. 1998). For example, a good conductor projects an image with a large centre region where the local EOD amplitude is increased, surrounded by a small rim area where the amplitude decreases compared to the amplitude in absence of an object. Images of non-conductors are of opposite arrangement: local EOD amplitude decreases centrally and slightly increases in the surrounding rim area (Fig. 2).

Electric images are always blurred, or "out of focus", since no focusing mechan-

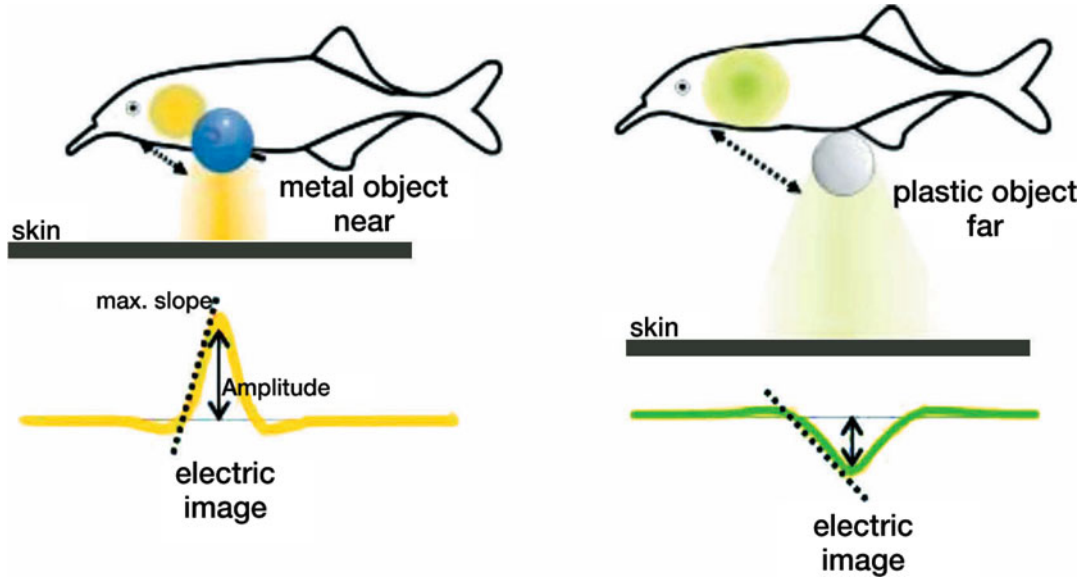


Fig. 2 The concept and properties of electric images in *G. petersii.* The 2-dimensional image depends on the distance of an object. With increasing distance (right), the location of the peak in the image remains constant, while the overall amplitude decreases, the image increases in width and the slope of the image (dotted lines) becomes shallower. Note that metal and plastic objects result in inverted images. The drawing was kindly provided by Dr. Jacob Engelmann

isms comparable to the lens of an eye exist. Hence when projecting 3-D objects on the 2-D sensory surface, there is no one-to-one relationship between spatial object properties and image shape. It follows that the fish has to use complex decoding algorithms in order to extract object information, such as size, shape and geometrical properties. In addition, electric images depend on object distance and location along the fish's body, the fish's body proportions, bending movements of the fish's body, the presence of additional objects, the background and many other factors. To make matters worse, electric images of two objects will fuse in a non-linear way depending on the distance between them (Caputi and Budelli 2006).

2.1 The production of EODs by electric organs

Weakly electric fishes produce electric impulses by muscle or nerve cell derived electric organs, which in the case of the African mormyridae lie in the caudal peduncle. Two basic types of EODs can be found: pulse type EODs, where the interval between two EODs is longer than the duration of a single EOD, and wave-type EODs, where discharges are produced directly one after another resulting in a sinusoidal wave signal (Fig. 3). In all cases, electric organ discharges are used for nocturnal orientation through active electrolocation and also for electro-communication (Moller 1995; Hopkins 2009). In both processes, the EOD waveform plays a critical role. The waveform of an EOD depends on the morphology of the electric organ and on the hormonal state of the animal. The electroreceptor organs involved in electrolocation are tuned to the characteristics of the self-produced EOD and thus can detect object-induced modifications of the local EOD. Most objects in the environment of the fishes are mainly resistive and thus only modulate the local EOD amplitude. However, animate objects also have capacitive properties, which lead to waveform and phase shifts of the local EOD in addition to amplitude changes (von der Emde and Bleckmann 1992).

For objects with complex impedances, amplitude and waveform modulations of

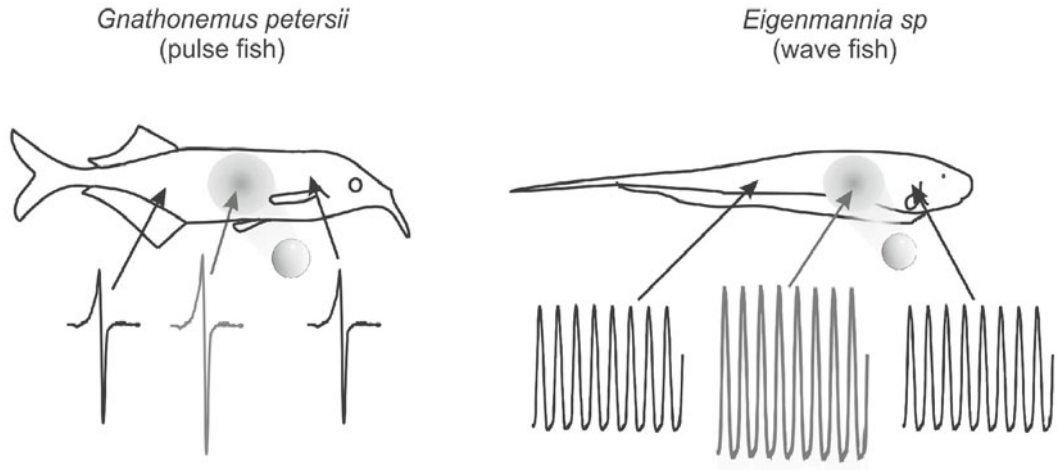


Fig. 3 Pulse-type (left) and wave-type EODs in *G. petersii* and *Eigenmannia* (right), respectively. In the centre of the electric image of a metal object (sphere), the amplitudes of the local EODs (grey) are increased relative to the amplitudes outside the image (black)

the EOD are frequency dependent. Pulse-type EODs contain many frequencies depending among other parameters on their duration. The range of capacitive object values an animal can detect thus depends on the duration of its EOD: shorter EODs enable the detection of lower capacitances and vice versa. Mormyrids have adapted their signals in such a way that the detectable range of capacitances corresponds to the range of capacitive values of animated objects found in the natural habitat of the fish (von der Emde and Ringer 1992).

Wavetype EODs also come in variety of waveforms and frequencies, which in different fish species span a range from below 100 Hz to more than 2000 Hz (Moller 1995). Until now, no study has clearly shown any correlation between EOD frequency and detection task in wave-fish. Instead, different EOD frequencies play an important role in electro-communication, for example, during species recognition, mate finding and for establishing dominance structures in groups of fishes.

2.2 Electroreceptor organs and pre-receptor mechanisms

2.2.1 Pre-receptor mechanisms

Before a sensory stimulus hits the receptor cells, it is modified by structures external to the receptors themselves. In active electroreception, such pre-receptor mechanisms have been studied first in some South American weakly electric fish, where they enhance the perception of electric images (Caputi et al. 2002; Pereira et al. 2005). Also in *G. petersii* pre-receptor mechanisms have been found that condition the sensory signals (Pusch et al. 2008). One of these is the fish's body which influences the electrical field through its shape and its low resistive internal electric properties. This allows a fish to focus the electric field on the tip of its Schnauzenorgan where the density of electroreceptors is highest. Electrical currents are additionally funnelled by the constantly open mouth ("funnelling-effect"), which leads to a homogenous voltage distribution in the nasal region (the region above the mouth) of the fish.

When *G. petersii* swims over the ground for foraging, they move their Schnauzenorgan rapidly, which might cause changes of the electric field distribution around the body and hinder active electrolocation with the Schnauzenorgan. Interestingly, however, it was demonstrated that the Schnauzenorgan takes the electrical field with it when it moves (Pusch et al. 2008). Thus, the tip-effect is still present also during movement, which results in constant stimulus condition at the tip of the Schnauzenorgan even during its rapid scanning movements during foraging. So, regardless of bending, the field is always optimal which helps the fish detect small objects that cause only minute amplitude changes at the Schnauzenorgan's tip.

2.2.2 Electroreceptor organs

All electroreceptive fishes possess so called ampullary electroreceptor organs which respond to low frequency electric signals of animate and inanimate origin. These receptor organs are used for passive electrolocation (Bodznick and Montgomery 2005; Wilkens and Hofmann 2005).

Weakly electric fishes from Africa and South America have, in addition, a second class of electroreceptor organs which are used to encode the amplitude and the phase of the fish's own EOD. These tuberous receptor organs are most sensitive to the dominant frequencies of the fish's own EOD. Like all mormyrid fish, *G. petersii* possess two types of tuberous electroreceptors, which are called Mormyromasts and Knollen-

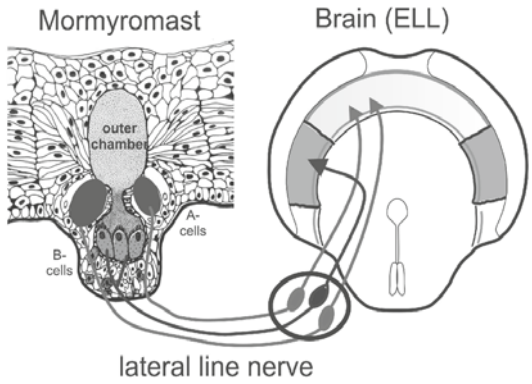


Fig. 4 A mormyromast electroreceptor organ (left) and its projection into the electrosensory lateral line lobe (ELL) in the brainstem (horizontal section on the right) of *G. petersii*. Electroreceptor cells (A- and B-cells) of the mormyromasts are innervated by lateral line nerve afferents, which project into somatotopic maps in two zones of the ELL

organs (Bennett 1965). Knollenorgans are exclusively used to detect the EODs of other mormyrids during electro-communication (Xu-Friedman and Hopkins 1999), while mormyromasts are responsive to the animal's own EOD and modulations of the EOD by the external environment (Bell 1990).

Each mormyromast houses two types of receptor cells that are tuned to different aspects of the signal-carrier, i. e. one channel for amplitude and one for waveform coding. Each mormyromast consists of two chambers containing several receptor cells. A-cells are found at the basal part of the outer chamber, while B-cells are located inside of an inner chamber (Fig. 4) (Bell et al. 1989). Both are innervated by separate nerve fibres, which project to the electrosensory lateral line lobe (ELL) in the brain, where type A and B afferent terminate in separate areas each forming a somatotopical map (Fig. 4) (Bacelo et al. 2008).

2.3 Behavior during active electrolocation

2.3.1 Electric signalling behavior

The pulse-type EODs of *G. petersii* are all-or-nothing events and their waveform cannot be modified by the animal on a short term basis (Hopkins 2009). However, the fish can change the temporal pattern of EODs produced and thus influence the number of EODs emitted within a certain time window. In wave-type EODs, this results in different frequencies of the wave signal, while in pulse-type EOD, the sequence of pulse intervals (SPI) changes. In pulse fish, this SPI functions in electro-communication by conveying different types of information between the fish (Carlson and Hopkins 2004). During active electrolocation, the SPI is important for regulating the flow of information about the environment to the animal.

In mormyrid pulse fish typical SPI can be observed, when electrically inspecting an object or during foraging. The important parameter during active electrolocation of objects seems to be a *regular pattern* of inter-pulse intervals, as suggested by several authors for various species of mormyrids (von der Emde 1992; Moller 1995; Carlson 2002;). All fish regularize their discharge activity during probing of an object. These regular patterns contrast with the variable discharge rates during swimming and also during food search. Regularization during object inspection may serve to keep receptors and associated brain structures on a constant level of adaptation, which may be especially important because a high degree of plasticity of electrosensory brain structures.

2.3.2 Locomotor behavior

When investigating a novel object, mormyrids perform probing motor acts (PMA), i. e. characteristic behaviours composed of a

series of swimming manoeuvres in close proximity to the object. Six types of PMA have been described, which all may serve to position the fish optimally for some aspect of active electrolocation (Toerring and Moller 1984; von der Emde 1992; Moller 1995). While all species of mormyrids perform similar PMAs, there exists one PMA which is only performed by *G. petersii*. During 'chin probing', *G. petersii* brings its movable chin appendix, the Schnauzenorgan, close to the object, almost touching it. The fish then moves the Schnauzenorgan over the object, following its contours. This behaviour resembles a haptic inspection of an object with the fingers of the human hand or the scanning movements of the fovea of the eye when looking at an object or inspecting a picture. The scanning of an object by these Schnauzenorgan movements provide fine detailed electrical information about the shape of the object (von der Emde and Fetz 2007; Engelmann et al. 2009).

When *G. petersii* searches for small insect larvae on the ground of the river, they never perform PMAs. Nevertheless, characteristic and stereotyped behavior occurs in these situations, helping to optimize sensory input about the prey. During foraging, *G. petersii* employs a characteristic swimming posture: they swim at a constant angle of their body axis of about 20° with their head towards the ground. With the tip of their Schnauzenorgan they almost touch the ground, moving it in a stereotyped fashion from left to right while swimming forward. During these very fast sweeping movements, the Schnauzenorgan scans a wide angle of the ground (Hollmann et al. 2008; von der Emde et al. 2008a).

When prey or another object of interest is encountered, the scanning movements of the Schnauzenorgan stop abruptly, and it is brought in a twitching movement towards the object for further exploration. In the case of prey, exploration is very brief and the fish tilts forward to suck up the insect larva. In order to acquire an object buried in the soil, the Schnauzenorgan is used as a burrowing stick to dig out the prey up to a depth of 2 or 3 centimetres.

The above described slanted swimming position during prey search also ensures that the nasal region, the skin area above the mouth and between the nares at the fish's head, is held fairly constant at an angle of about 50° relative to the ground. It thus points forward and slightly upward so as to be in an optimal position to detect approaching objects such as obstacles. When the fish approaches an obstacle, the object will cause an electric image to be projected onto the nasal region and will thus be detected and identified. The nasal region, which contains an exceptionally high density of electroreceptors, is thus used like the fovea in the retina of the eye.

2.4 Perception of objects by weakly electric fish

During active electrolocation weakly electric fish can recognize several object properties. Among these is the electrical resistance of an object, which is determined by measuring the amplitude changes of the locally perceived EOD caused by the object (Lissmann and Machin 1958). Capacitive object properties, i. e. an object's complex electrical impedance, are perceived by measuring the local EOD waveform distortions with the A- and B-type electroreceptor cells within each mormyromast organ (von der Emde and Bleckmann 1992).

Mormyrids can also localize objects in 3-dimensional space during active electrolocation (von der Emde et al. 1998). This ability is based only on distance and is independent of the size or electrical properties of the object. *G. petersii* thus has a true sense of depth perception.

If an object moves away from the fish

three things happen: the size of the electric image gets larger, the amplitude in the image's centre decreases and the image gets more fuzzy, i. e. the rise (or fall) of the amplitude towards the centre is less steep. The latter parameter, called the *slope* of the image, together with the amplitude change in the centre turned out to be crucial for distance measurements by *G. petersii:* The animals measure the *slope/amplitude ratio* of the electric image, which depends only on object distance and not on size or other object properties (Fig. 2) (Schwarz and von der Emde 2001).

Even though the electric sense lacks focusing mechanisms, weakly electric fish can also perceive an object's three-dimensional shape. In a two-alternative forced-choice procedure *G. petersii* quickly learned to recognize objects of various shapes and to discriminate them from objects of other shapes (von der Emde and Fetz 2007). Shape recognition persisted even when the objects were rotated in space, indicating a viewpoint independent recognition of objects. In additional experiments, *G. petersii* demonstrated size constancy during object recognition (von der Emde and Fetz 2007). For analyzing the shape or size of an unknown object, fish have to perform probing motor acts, i. e. they have to swim around the object scanning it with their sensory surface from several viewpoints. This is in contrast to distance measurements, which can be achieved instantly and do not require scanning movements.

As mentioned above, the electric images of objects located close to each other near the fish will fuse in a nonlinear way leading to complex electric images. In spite of this effect, *G. petersii* is able to perceive the shape of an object even when it is positioned right in front of a large background (von der Emde et al. 2008a). When trained to discriminate between two objects placed next to each other and a solid object of the same length, *G. petersii* was able to detect gaps as small as 2–3 mm between the two objects (K. Behr, unpublished). At present it is not known what physiological mechanisms the fish use to perform these tasks.

3. Developing technical electrolocation sensors

Given the remarkable abilities of weakly electric fish to detect and analyze objects during active electrolocation, the idea was pursued of using the same principles as the fish in technical sensor applications (Schwarz et al. 2001). A sensor system working according to these principles might benefit from the fact that many of the mechanisms of active electrolocation in weakly electric fish are well understood. When designing such a system, it is not the goal to build an exact copy of an electric fish or to design a sensor that looks like these animals. Instead, a Bionik-approach was chosen, i. e. our goal was to confer the *principles* of active electrolocation to sensor systems optimized to solve specific remote sensing tasks (von der Emde et al. 2008b).

3.1 Mathematical modelling

When trying to understand the working principles of a complex system, one is faced with a great number of variables and possible solutions. Under these conditions it is helpful to develop a mathematical model of the system in order to be able to quickly test and try out a large number of parameter values and possible solutions. Such a model should replicate the findings of the biological experiments carried out with the natural model, i. e. with weakly electric fish.

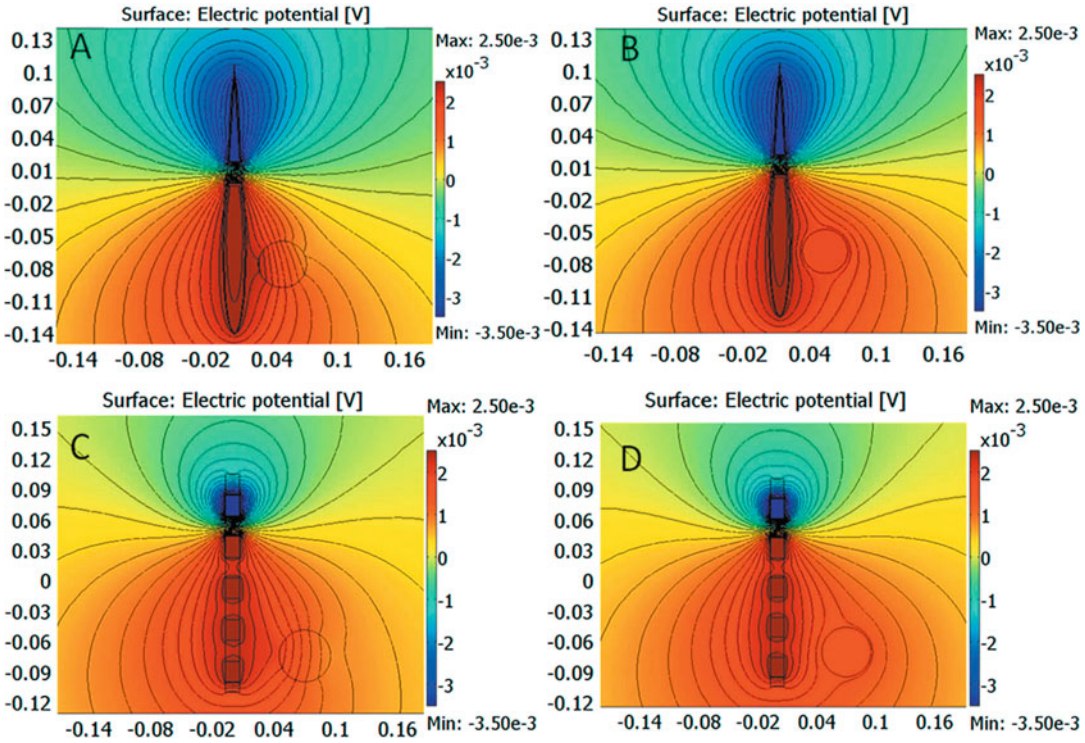


Fig. 5 Modelling results of the electrical field produced by an electric fish (**A**, **B**) and a rod-like technical sensor system (**C**, **D**) at the head-positive phase of the EOD. Fields are colour coded and additionally depicted by equipotential lines. Equipotential lines are denser near the skin of the fish when a target of brass (conductor) is present close to the skin (**B**, **D**), leading to increased amplitudes of the electric signal. In contrast, equipotential lines pass through an insulating target object (**A**, **C**). Similar behaviour of equipotential lines can be seen when the target is placed near the technical sensor model or the fish model

After these biological systems have been modelled accurately, the next step will be to apply the model to a particular sensor task.

During active electrolocation, a nearby object differing in its electrical impedance from the surrounding water will change the flow of the electric current produced by the fish. These perturbations give rise to changes in the local potential differences across the skin of the fish (transdermal voltage) at those skin regions facing the object. In order to understand the range of transdermal voltage changes detected by the fish, we first modelled the electrical field of a weakly electric fish based on previous modelling studies (Rasnow 1996; Budelli and Caputi 2000; Kelly et al. 2008;). Our models were designed in COMSOL Multiphysics (www.comsol.com). The electric field around the fish was calculated using Poisson's equation:

$$\nabla\Phi = \frac{\rho_0}{\varepsilon_0 \varepsilon_r} \quad (1)$$

where ϕ is the potential in volts, ρ_0 is the charge density in C/m^3 and the permittivity of free space (ε_0) is taken to be 8.854 × 10–12 F/m. The conducting medium was assumed to be water with a conductivity of 400 µS/cm and a relative permittivity (ε_r) equalling 80. Because tuberous electroreceptor organs of weakly electric fish respond to a frequency range between 1 and 5 kHz,

the frequency used in our model was fixed at 1 kHz. The fish morphology and the electric field parameters for the model were taken from the existing literature and our own results (von der Emde et al. 2008b).

Figure 5A, B show the modelling results for the electrical field surrounding a weakly electric fish with a length of 20 cm in the presence of a non-conducting (conductivity = 0 $S*m^{-1}$, A) or a conducting spherical object (conductivity = 3.5×10^7 $S*m^{-1}$, B). Field voltages are shown in false colours with positive voltages in red and negative voltages in blue. The effects of the two types of objects are clearly opposite. The conducting (metal) object has no electrical field gradient inside itself and the surface of the target is equipotential. For insulating objects, the potential lines continue inside the target, which results in a potential difference along the surface. These differential effects will influence the electrical current distribution on the skin of the animal in an opposite way, leading to a differential response of the electroreceptor organs embedded in the skin. Our modelling results thus reflect accurately the results obtained in biological experiments that measured the electric images cast by real objects onto the skin of weakly electric fishes (e. g. in Pusch et al. 2008).

In order to model a technical sensor system, the geometry of the fish model was modified in such a way that a non-conducting rod-like structure of the artificial sensor resulted. The electric source (analogous to the electric organ of the fish) was assumed to be an electrode at the end of the rod with a negative current density. The electroreceptors were assumed to be electrodes with positive current densities. The initial conditions were calculated in the absence of an electrolocation target. The voltage was noted along the respective tangential axes of the artificial sensor, again in analogy to the fish. In order to measure the effective changes caused by the presence of a target, the initial conditions were subtracted. As in the fish model, the targets consisted of a metal or a plastic sphere. The results of our simulations are shown in Fig. 5C and D. In addition, we modelled the effects of moving the target object away from the fish or the rod up to a distance of 10 cm in intervals of 1 cm. As the target moved away from the central axis, the change in potential decreased both in the fish and the technical model in a similar way (K. Mayekar, unpublished observations).

3.2 Measurements

In addition to modelling biological and technical sensor systems, we initially started to design prototypes of technical sensors, which can be used for physical measurements. The first prototypes consisted of three parts: an array of electrodes, an analog device for signal generation and processing, and a computer (Fig. 6). In later versions, the analog device was partly replaced and incorporated into the logical computer device. The electrodes build up an electrical field in the medium and simultaneously measure it at several points. In order to do so, the electrodes were usually arranged in lines with fixed distances between them. One or two outer electrodes formed one pole of the emitting system; all the other ones (between 1 or 16 electrodes) formed the second pole. Analogous to wave-type electric fish, the electrical signal that was emitted by the system was a sine wave. The voltage between the poles was amplified and filtered, analog to digital converted and transferred to the computer for further analysis. In order to detect, localize or analyze an object near the sensors, the differences between the undisturbed field without an object present and the field distorted by the object was calculated and processed. Thus, at the beginning of each session the sensor had to obtain at

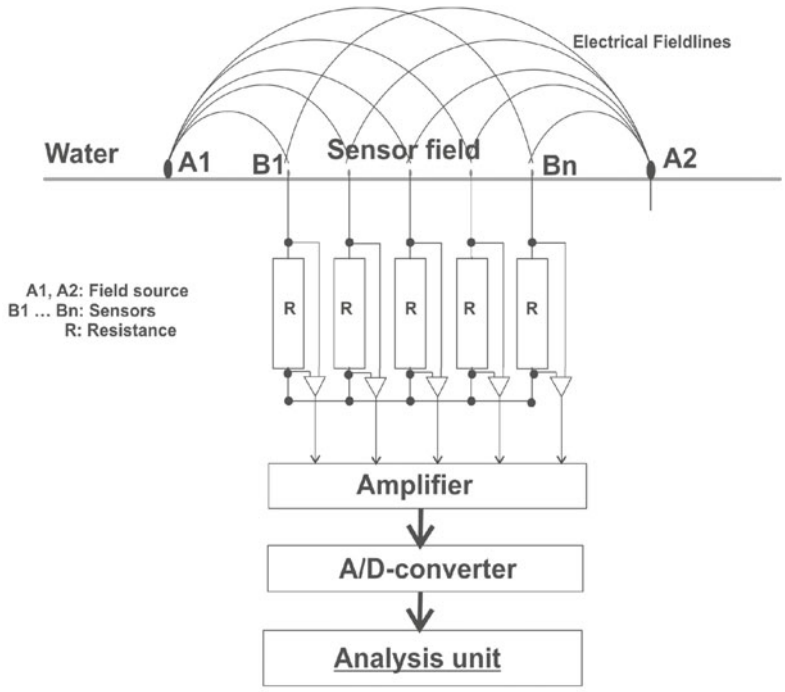


Fig. 6 Schematic set-up of a technical active electrolocation sensor. An electrical field is build up by one of the A-electrodes (both fields are shown here). Electrodes B_1 ... B_n form an array of measuring electrodes. Each B-electrode measures the local voltage, which might be changed by the presence of an object within the electrical field

least one measurement in the absence of the object in order to obtain a baseline level of voltages, which could later be used for comparisons. In all following object investigations, only the difference between this first calibration and the following measurements was used.

Different types of sensors that work according to the principles of active electrolocation can be used for a wide range of industrial applications. The requirements for the different technical purposes are variable, requiring the development of an optimized sensor system for each task.

One of the big advantages of technical electrolocation sensors is that they can work under difficult environmental conditions, such as under high pressure or in high or even changing temperatures. In addition, these sensors can work in polluted environments, e. g. in murky waters, or in media containing contaminations.

3.2.1 Target detection

One of the easiest tasks of technical electrolocation sensors is to detect the presence or the disappearance of an object. A change of sensor output above a certain predefined threshold can be used for such a task. Continuous measurements of sensor input and comparisons of the inputs over time will result in a high sensitive detection system for objects of different materials. Because the sensors can detect both electrical conducting objects (metal) and non-conductors (plastic) they are flexible systems, with a single sensor being able to detect the presence of most kinds of objects (Fig. 7).

3.2.2 Remote distance measurement

An interesting application of an active electrolocation sensor is the accurate determination of the distance to an object at short range. One of many possible applications of

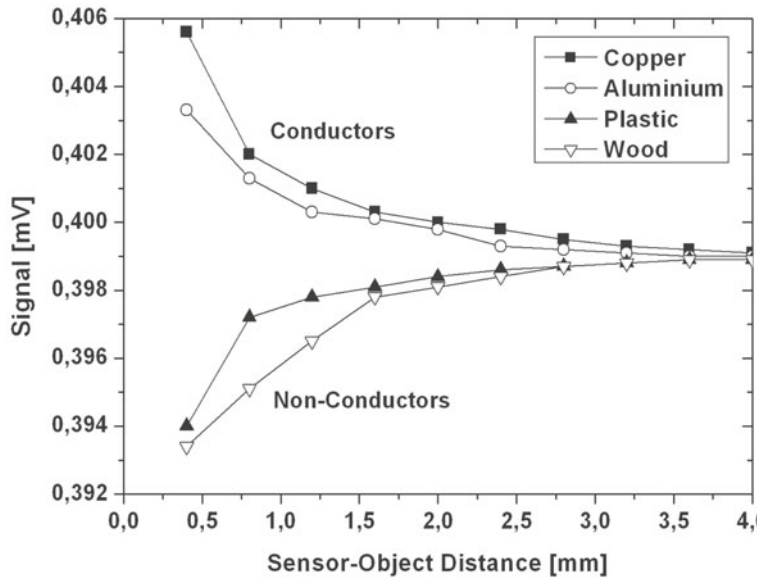


Fig. 7 Distance measurements with a technical, 2-pole active electrolocation sensor. The voltage measured by the sensor is plotted versus the distance of the sensor from the target object. At large distances, voltage amplitudes are medium. Voltages increase (conductors) or decrease (non-conductors) when the sensor-object distance is reduced. Note that the sensor not only responds to distance changes but also to different types of materials

such systems could be in industrial production, where it is important to know the exact position and the distance of a piece of work to a machine or parts of a machine. To solve such a problem, an initial prototype sensor was built consisting of only two electrodes. These were used to build up an electric field in the medium and to simultaneously measure this field in the presence or absence of objects. A prerequisite for the sensor to work under these conditions is to calibrate the system in the absence of the object. After calibration, the base line signal is stored and during subsequent measurements it is subtracted from the current signal. If the object is known and has been measured before, the resulting difference can be used to calculate the distance to the object. Figure 7 shows the results of measurements with a two-electrode electrolocation sensor measuring the distance of four types of objects. Up to a distance to about 2.5 cm a specific signal could be measured for each object, which might be used to determine object distance when the type of object is known.

In some industrial applications, it is either impossible to calibrate the sensor before each measurement, or the physical properties of the medium and the object are not constant. To cope with these conditions, we designed another sensor-system consisting of more than two electrodes. With this type of multi-electrode sensor it is possible to apply the algorithms previously found to be used by weakly electric fish when measuring distances to objects, i. e. using the above mentioned slope/amplitude ratio (von der Emde et al. 1998). It turned out that using this method it was possible to roughly determine distances of a variety of objects, independent of object properties such as size or material (S. Schwarz, unpublished observation).

3.2.3 Detection of material properties

The ability of weakly electric fish to recognize the electrical properties of objects (and thereby their material properties) can be assigned to technical sensor systems working according to the principles of active electrolocation (Fig. 7). By designing a sensor system that scans an object and measures

these properties it may be possible to detect impurities of the material composition of an object. Moreover, mechanical inhomogeneities inside an object, such as cracks or holes, can also lead to changes in current flow through an object and thus might be detectable by an active electrolocation sensor.

Active electrolocation sensors could also be applied in medical diagnostics and provide imaging information in media which are difficult to access using other systems. One example for this may be the detection and recognition of "vulnerable plaques" in blood vessels, which are one major cause of heart attacks. We are developing a system using a catheter based microelectrode array sensor, which, like the electric fish, emits an electrical signal and simultaneously detects the three dimensional electrical field changes caused by the blood vessel wall. Preliminary experiments revealed that such a sensor system can distinguish biological materials with different dielectric properties (Metzen et al. 2010). We therefore assume that a catheter sensor based on active electrolocation will be able to detect and localize different types of "plaques" in blood vessels. In addition, such a system will be able to determine the basic composition and size of such plaques and therefore might constitute a future diagnostic tool for the prevention of heart attacks.

References

Bacelo J, Engelmann J, Hollmann M, von der Emde G, Grant K (2008) Functional foveae in an electrosensory system. J Comp Neurol 511: 342–359

Bell CC (1990) Mormyromast electroreceptor organs and their afferent fibers in mormyrid fish. III. Physiological differences between two morphological types of fibers. J Neurophysiol 63: 319–332

Bell CC, Zakon H, Finger TE (1989) Mormyromast electroreceptor organs and their afferent fibers in mormyrid fish: I. Morphology. J Comp Neurol 286: 391–407

Bennett MVL (1965) Electroreceptors in mormyrids. Cold Spring Harbor Symp Quant Biol 30: 245–262

Bodznick D, Montgomery JC (2005) The Physiology of low-frequency electrosensory systems. In: Bullock TH, Hopkins CD, Popper AN, Fay RR (eds) Electroreception. Springer, New York, pp 132–153

Budelli R, Caputi AA (2000) The electric image in weakly electric fish: perception of objects of complex impedance. J Exp Biol 203: 481–492

Bullock TH, Northcutt RG, Bodznick DA (1982) Evolution of electroreception. Trends Neurosci 5: 50–53

Caputi AA, Budelli R (2006) Peripheral electrosensory imaging by weakly electric fish. J Comp Physiol A 192: 587–600

Caputi AA, Budelli R, Grant K, Bell CC (1998) The electric image in weakly electric fish: Physical images of resistive objects in *Gnathonemus petersii*. J Exp Biol 201: 2115–2128

Caputi AA, Castelló ME, Aguilera PA, Trujillo-Cenoz O (2002) Electrolocation and electrocommunication in pulse gymnotids: signal carriers, prereceptor mechanisms and the electrosensory mosaic. J Physiol (Paris) 96: 493–505

Carlson BA (2002) Electric signalling behavior and the mechanisms of electric organ discharge production in mormyrid fish. J Physiol Paris 96: 405–419

Carlson BA, Hopkins CD (2004) Stereotyped temporal patterns in electrical communication. Anim Behav 68: 867–878

Engelmann J, Nöbel S, Röver T, von der Emde G (2009) The Schnauzenorgan-response of *Gnathonemus petersii*. Frontiers in Zoology 6 21

Hollmann M, Engelmann J, von der Emde G (2008) Distribution, density and morphology of electroreceptor organs in mormyrid weakly electric fish: anatomical investigations of a receptor mosaic. J Zool 276: 149–158

Hopkins CD (2009) Electrical perception and communication. In: Squire L (ed) Encyclopedia of neuroscience. Academic Press, Oxford, pp 813–831

Kelly M, Babineau D, Longtin A, Lewis JE (2008) Electric field interactions in pairs of electric fish: modeling and mimicking naturalistic inputs. Biol Cybern 98: 479–490

Lissmann HW, Machin KE (1958) The mechanism

of object location in *Gymnarchus niloticus* and similar fish. J Exp Biol 35: 451–486

Metzen M, Biswas S, Bousack H, Gottwald MG, Mayekar K, von der Emde G (2010) A biomimetic active electrolocation sensor for detection of atherosclerotic lesions in blood vessels. IEEE Sensors Journal, DOI: 10.1109/JSEN.2010.2079928

Moller P (1995) Electric fishes. History and behavior. Chapman & Hall, London

Pereira ASF, Centurión V, Caputi AA (2005) Contextual effects of small environments on the electric images of objects and their brain evoked responses in weakly electric fish. J Exp Biol 208: 961–972

Pusch R, von der Emde G, Hollmann M, Bacelo J, Nöbel S, Grant K, Engelmann J (2008) Active sensing in a mormyrid fish: electric images and peripheral modifications of the signal carrier give evidence of dual foveation. J Exp Biol 211: 921–934

Rasnow B (1996) The effects of simple objects on the electric field of *Apteronotus*. J Comp Physiol A 178: 397–411

Schwarz S, Hofmann MH, von der Emde G (2001) Weakly electric fish as a natural model for industrial sensors. BIONA-Report 15: 142–157

Schwarz S, von der Emde G (2001) Distance discrimination during active electrolocation in the weakly electric fish *Gnathonemus petersii*. J Comp Physiol A 186: 1185–1197

Toerring MJ, Moller P (1984) Locomotor and electric displays associated with electrolocation during exploratory behavior in Mormyrid fish. Behav Brain Res 12: 291–306

von der Emde G (1992) Electrolocation of capacitive objects in four species of pulse-type weakly electric fish. II. Electric signalling behavior. Ethology 92: 177–192

von der Emde G (1998) Electroreception. In: Evans DH (ed) The physiology of fishes. CRC Press, Boca Raton, Florida, pp 313–343

von der Emde G, Amey M, Engelmann J, Fetz S, Folde C, Hollmann M, Metzen M, Pusch R (2008a) Active Electrolocation in *Gnathonemus petersii*: Behaviour, sensory performance, and receptor systems. J Physiol (Paris) 102: 279–290

von der Emde G, Bleckmann H (1992) Differential responses of two types of electroreceptive afferents to signal distortions may permit capacitance measurement in a weakly electric fish, *Gnathonemus petersii*. J Comp Physiol A 171: 683–694

von der Emde G, Bleckmann H (1998) Finding food: Senses involved in foraging for insect larvae in the electric fish, *Gnathonemus petersii*. J Exp Biol 201: 969–980

von der Emde G, Bousack H, Huck C, Mayekar K, Pabst M, Zhang Y (2008b) Electric fish as natural models for technical sensor systems In: Rodríguez-Vázquez AB (ed) Bioengineered and bioinspired systems, Dresden

von der Emde G, Fetz S (2007) Distance, shape and more: recognition of object features during active electrolocation in a weakly electric fish. J Exp Biol 210: 3082–3095

von der Emde G, Ringer T (1992) Electrolocation of capacitive objects in four species of pulse-type weakly electric fish. I. Discrimination performance. Ethology 91: 326–338

von der Emde G, Schwarz S, Gomez L, Budelli R, Grant K (1998) Electric fish measure distance in the dark. Nature 395: 890–894

Wilkens L, Hofmann M (2005) Behavior of animals with passive, low-frequency electrosensory systems. In: Bullock T, Hopkins C, Popper A, Fay R (eds) Electroreception. Springer, pp 229–263

Xu-Friedman MA, Hopkins CD (1999) Central mechanisms of temporal analysis in the knollenorgan pathway of mormyrid electric fish. J Exp Biol 202: 1311–1318

Microsecond and millisecond time processing in weakly electric fishes

22

Masashi Kawasaki

Contents

Abstract

Weakly electric fishes possess two distinctive electrosensory systems that are specialized to detect amplitude and time (phase) information. The amplitude sensitive system detects electric feedback signals from the fish's own electric organ discharges that are altered by the resistive component of nearby objects. The time (phase) sensitive system, on the other hand, detects signal times that are altered by the capacitive component of the objects. These subcomponents of the electrosensory system are fundamental to the fish's ability to obtain the electric image of its immediate surroundings. The jamming avoidance response is a behavior in which the fish alter the frequency of their electric organ discharges to avoid mutual jamming of the electrolocation behavior. The information necessary to perform correct jamming avoidance responses is encoded in the time pattern of amplitude and time signals that occurs in the millisecond time scale. The time signal is encoded in the difference in phases between signals at different locations on the body surface. This differential phase sensitive system operates on the microsecond time scale. Our behavioral experiments demonstrated that even submicrosecond phase differences could be detected by the electrosensory system and manifested in the jamming avoidance responses. Neurons of the central nervous system that are specialized to process amplitude and differential phase information were found in the electrosensory lateral line lobe. In the midbrain, neurons are sensitive to specific time patterns of activities of the amplitude and differential phase sensitive neurons projecting from the lower station. These midbrain neurons examine time pattern of synaptic potentials and exhibit responses necessary for the jamming avoidance response. The electrosensory system of electric fishes is one of the best-understood systems for temporal coding of sensory information.

Masashi Kawasaki
University of Virginia, Department of Biology
Gilmer Hall, McCormick Road
Charlottesville, VA 22904, USA
e-mail: mk3u@virginia.edu

1. Introduction

Wave-type electric fishes emit sinusoidal electric organ discharges (EODs) at several hundred Hz from their tail setting up an alternating (AC) electric field around the body. Electric field distortion by nearby objects is sensed by electroreceptors in the skin allowing the fish to detect the location, size, shape, and the electrical properties of the object. In this electrolocation behavior, fish are capable of detecting resistive as well as capacitive components of the objects (Fig. 1A) (von der Emde 1992) (see also Chapter 21 by G. von der Emde). Resistive and capacitive components are sensed by amplitude and phase sensitive electrosensory systems, respectively.

The frequency of the EOD is highly stable (Bullock 1970; Guo and Kawasaki 1997). When two fish with similar EOD frequencies meet, however, their electric fields interfere and their ability to electrolocate deteriorates (Heiligenberg 1975). To avoid such mutual jamming, the fish perform the jamming avoidance responses (JARs) in which they shift their discharge frequencies so as to increase their frequency difference (Bullock et al. 1975; Kawasaki 1993) (Fig. 1B). The fish is capable of determining *the sign* of the frequency difference (Df) between its own discharge and that of a neighbor and performs a frequency-increasing or frequency-decreasing JAR appropriately. When a fish encounters a neighbor with a discharge frequency lower than its own, it increases its frequency; when it encounters a neighbor with a higher discharge frequency, it decreases its own.

Detailed behavioral studies identified a computational algorithm by which the central nervous system determines the sign of Df in *Eigenmannia* and *Gymnarchus* (Heiligenberg 1991; Kawasaki 1993). During the JAR, electroreceptors are exposed to the mixture of the fish's own EODs and those of its neighbor. The EOD mixture contains two elements of temporal code that jointly carry information as to the sign of Df. The temporal coding occurs in microsecond and millisecond time scales. This article presents neuronal mechanisms for the processing of

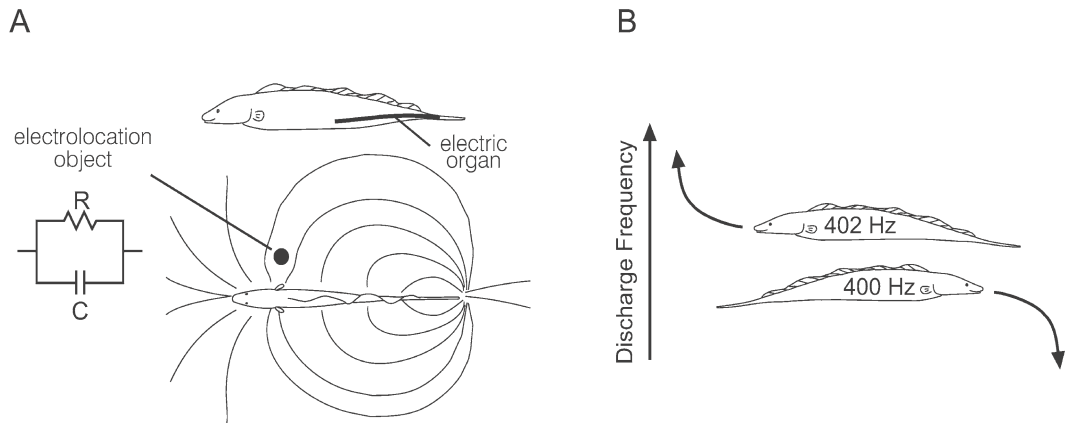


Fig. 1 Electrical behaviors in weakly electric fish. **A** Electrolocation. Objects with electrical resistivity or capacity different from those of the surrounding water cause a distortion of the electric field set by the electric organ of the fish. The distortion is monitored by electroreceptors in the skin for electrolocation of the object. **B** The jamming avoidance response. Fish change their otherwise constant discharge frequencies away from each other to avoid mutual jamming of their electrosensory systems

time in *Gymnarchus* and compares microsecond time processing in different species.

Sensing mechanisms in the microsecond time scale is widely found in the electrosensory system of the electric fishes and in the auditory systems in vertebrate and invertebrate animals. A hearing aid for the humans designed after a microsecond sensing auditory organ of an insect is an outstanding example of biologically inspired engineering product (Miles and Hoy 2006).

2. Neuronal mechanisms for microsecond time sensing in *Gymnarchus*

The first temporal cue for the JAR is phase difference. Due to differential contamination of the fish's own EODs by the neighbor's EODs across body areas during the JAR, different areas of the body experience phase modulations of different depths (Kawasaki 1993). The phase difference occurs in the microsecond time scale. Even smaller phase differences (100 to 300 nanoseconds) still induce JARs (Fig. 2) (Guo and Kawasaki 1997).

The precise phase (zerocrossing time) of each cycle of the electrosensory signal at different body areas is conveyed by phase-locked action potentials of the time coding afferent, the S-type afferent (Fig. 3) (Kawasaki and Guo 1996). The afferent fibers synapse onto the soma of the giant cells that project to the inner cell layer (ICL) of the medial zone of the electrosensory lateral line lobe (ELL). Both S-type afferents and giant cells fire one action potential in response to one zerocrossing of the sensory signal with a precision of a few tens of microseconds. The afferent fibers bifurcate before they terminate onto the giant cell soma and their branches also terminate onto the inner cell layer (Fig. 4). Thus the terminals of giant cells and afferents converge at the inner cell layer. By individually labeling component neurons, Matsushita and Kawasaki (2004) described detailed cytoarchitecture of this convergent

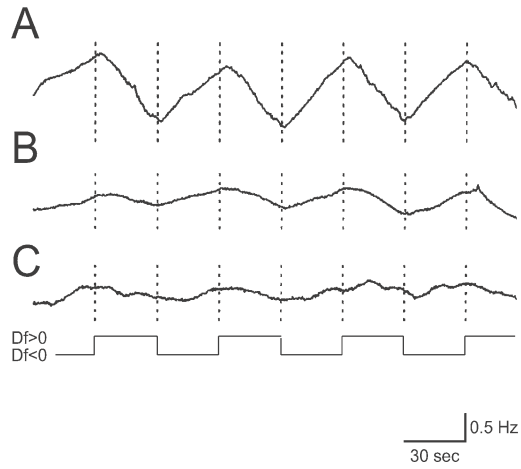


Fig. 2 Jamming avoidance responses to small amplitude and phase differences. JARs of rising and falling frequency, respectively, were induced alternately every 30 seconds (dotted lines). The bottom trace indicates the sign of the frequency change (Df). Magnitudes in amplitude/differential phase modulations were 0.58 %/2.8 µsec in **A**, 0.058 %/280nsec in **B**, and 0.0192 %/92.6nsec in **C**. Modified from Guo and Kawasaki (1997)

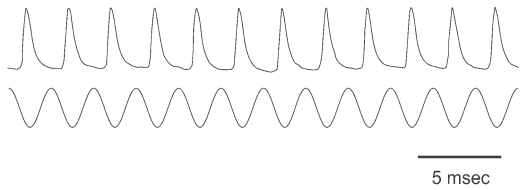


Fig. 3 Phase-locked responses of the S-type afferent fiber in *Gymnarchus* (**top**) in response to sinusoidal sensory stimulation (**bottom**). Each cycle of the electrosensory stimulus evokes a phase-locked action potential with precision. The giant cells also respond in a phase-locked manner with slightly longer latency

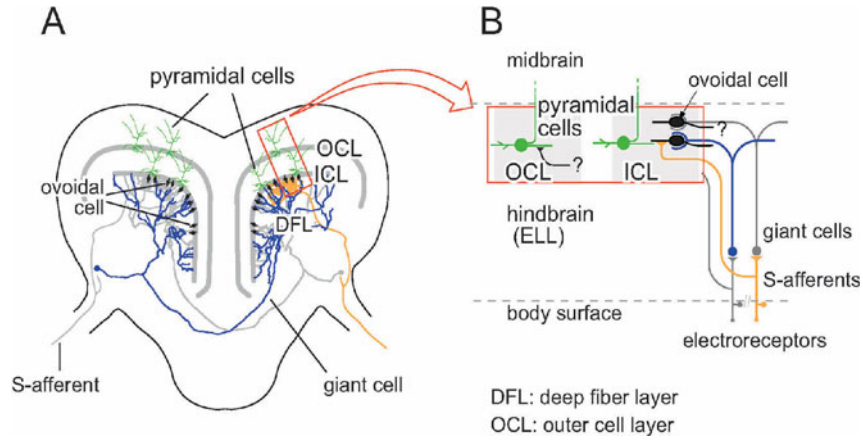


Fig. 4 The phase comparison circuit in the electrosensory lateral line lobe (ELL) of *Gymnarchus*. **A** Transverse view of the ELL. The S-type afferents terminate onto the soma of giant cells which terminate in the inner cell layer (ICL) – a layer in the ELL. A branch of afferents also terminates in the ICL. **B** The area of an ovoidal cell in A is expanded. The ovoidal cells are embraced by a large terminal of a giant cell and an afferent fiber. The pyramidal cells (green) are output neurons of the phase comparison circuit. The pyramidal cells respond to phase differences by changing the spike rate. Modified from Matsushita and Kawasaki (2004)

site in the inner cell layer by light and electron microscopy (Fig. 5). They found that the ovoidal cell is crucially located in the circuit for phase comparison. Individual terminals of the giant cell embrace ~85 % of the surface area of the soma of an ovoidal cell with an extraordinarily large synapse. Only chemical transmission seems to take place at this giant synapse. The dendrite of the ovoidal cell, which extends out of the giant synapse, is embraced with a single large terminal of an S-afferent fiber. Electron microscopic observations have suggested both chemical and electrical transmission at this synapse (Matsushita and Kawasaki 2004). Thus, each ovoidal cell receives one synapse from a giant cell and one synapse from an S-afferent. This unusual morphology strongly suggests that the ovoidal cell is the phase comparator and detects the phase difference between the terminals of a giant cell and an afferent.

Although to record the electrical activity of the ovoidal cells has not been successful, we have recorded from the output

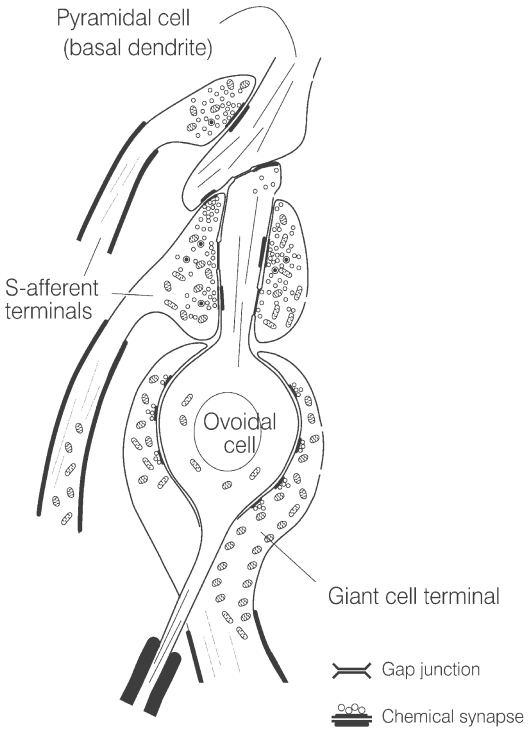


Fig. 5 The ovoidal cell and its two input synapses in the ELL of *Gymnarchus*. The projection sites of ovoidal cell axons are unknown. The dendro-dendritic connection to the pyramidal cell is thought to be the output path. From Matsushita and Kawasaki (2004)

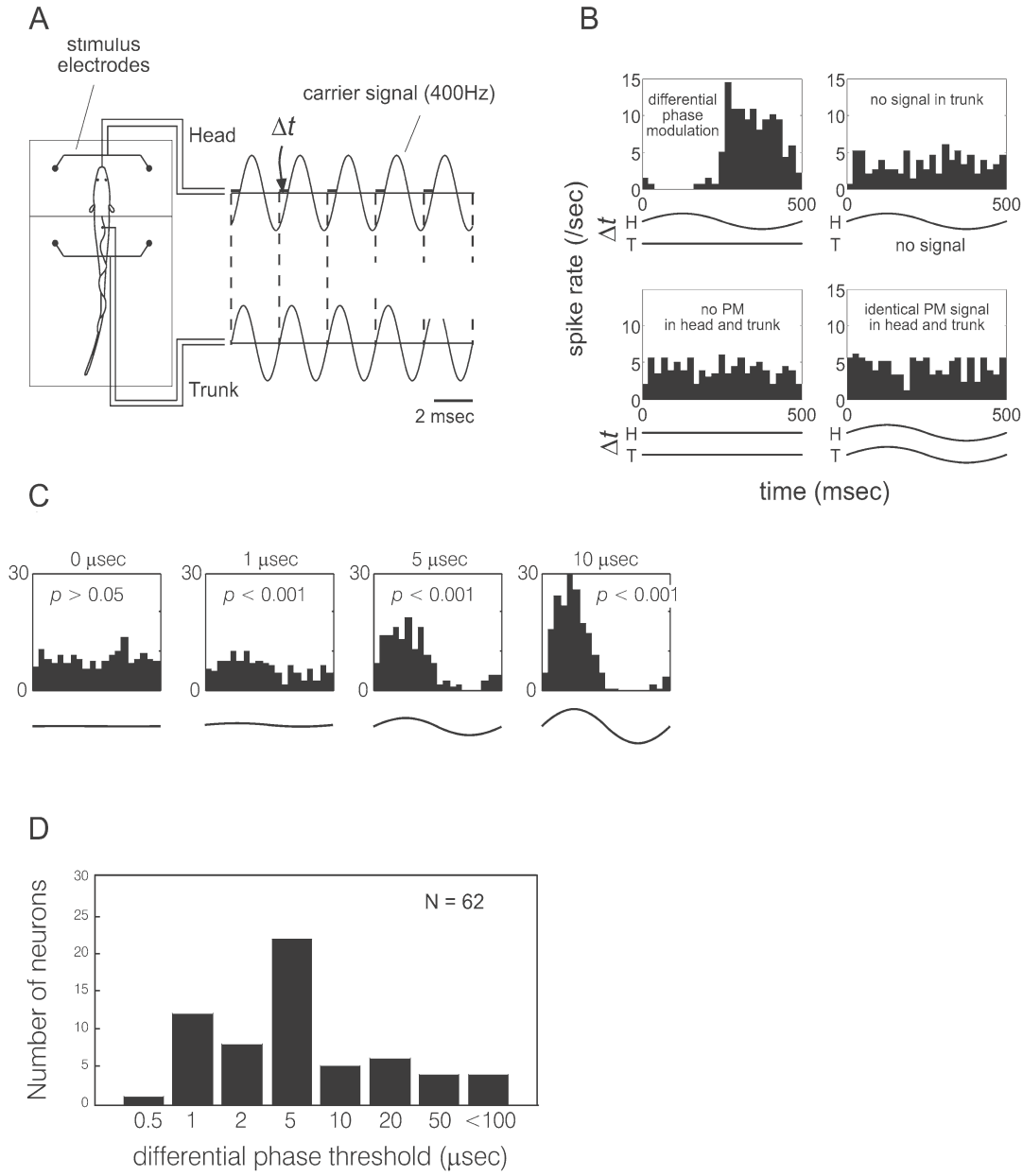


Fig. 6 Responses of the pyramidal cells in the electrosensory lateral line lobe of *Gymnarchus* to phase differences. **A** The 'phase chamber' in which the head and trunk portions of the body were electrically isolated and given separate sinusoidal signals with precisely controlled phase. **B** An example neuron showing the responsiveness to phase differences between the chambers but not to changes in phase in individual chambers. **C** An example of threshold measurement. This neuron showed statistically significant responses to phase modulation depths larger than 1 μsec. **D** Distribution of thresholds in 62 pyramidal cells. Modified from Matsushita and Kawasaki (2005)

neuron of the circuit, the pyramidal cells, and have shown their sensitivity to phase differences (Matsushita and Kawasaki 2005). The curarized preparation was placed in a 'phase chamber' in which different parts of the body surface were given phase controlled sensory signals (Fig. 6A). The pyramidal cells responded to *differences* between the phases of two signals but not to phase modulation in one chamber alone (Fig. 6B). Threshold values for the magnitude of detectable phase differences were on the order of a few microseconds (Fig. 6C, D).

3. Neuronal mechanisms for millisecond time sensing in *Gymnarchus*

In addition to phase difference information, wave-type electric fishes use amplitude cues for the JAR. During the JAR, electroreceptors are exposed to electrosensory signals from the fish's own and a neighbor's EODs. The signal mixture exhibits modulations in amplitude as well as in phase difference. As shown in Fig. 7, the time relations between amplitude and differential phase modulations are unique for Df < 0 and Df > 0, or JARs consisting of an increase or a decrease of discharge frequency. Behavioral experiments demonstrated that the fish detects the sense of rotation by examining the temporal sequence of four sensory events occurring in the amplitude and time difference axes, amplitude increase and decrease, phase advance and delay (Kawasaki 1993). The Lissajous graphs in Fig. 7 complete a rotation every 500 milliseconds when |Df| = 2 Hz. The increase and decrease of amplitude are represented by excitatory and inhibitory amplitude sensitive neurons in the electrosensory lateral line lobe (ELL) in the hind brain that fire a burst of action potentials at the rising and falling epochs of the amplitude modulations, respectively. Likewise, there are two types of neurons in the ELL that are sensitive to phase differences and fire a burst of action potentials during phase difference advances and delays, respectively. These ELL neurons respond identically to either sense of rotation, because they respond only to one parameter, amplitude or time difference. Only the temporal relation between these par-

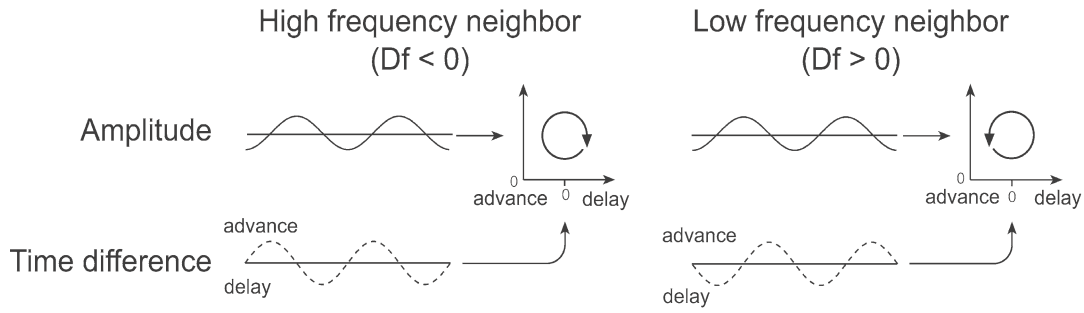


Fig. 7 Stimuli causing the JAR in wave-type electric fishes. Left and right panels, respectively, depict situations where the fish decreases (Df < 0) or increases (Df > 0) its firing frequency of the EOD in the JAR. Due to contamination of the fish's own EOD by a neighbor's EOD, modulations occur in both amplitude and time difference between afferent signals from different body areas. The strongest JAR is induced when the frequency difference between the two fish is 2 Hz in most individuals of *Gymnarchus*. The Lissajous graph completes a cycle in 0.5 sec when the frequency difference is 2 Hz

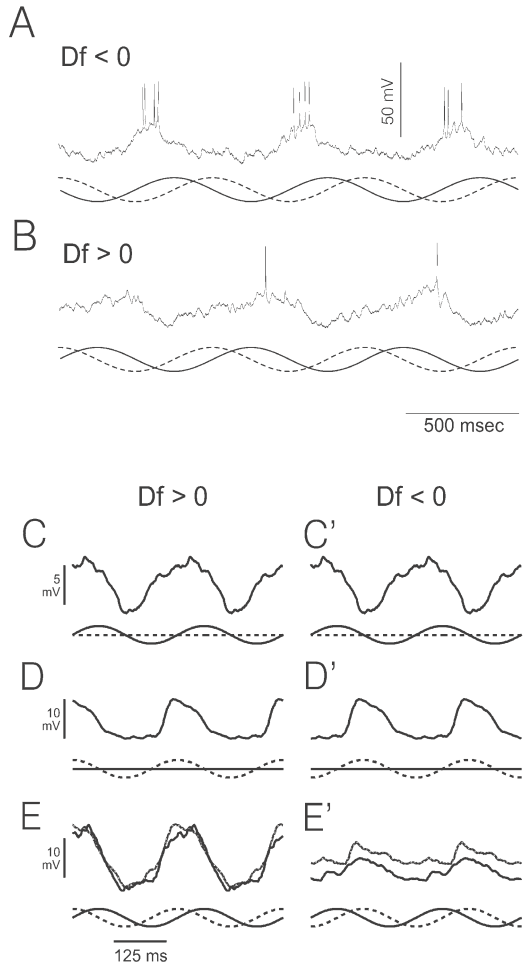


Fig. 8 Intracellular recordings from a sign-selective neuron in the torus semicircularis of *Gymnarchus*. **A** and **B** An example neuron showing a spike preference to Df < 0. **C** and **D** Spike-removed membrane potentials in response to singular presentation of amplitude modulation (solid line) or modulation in time difference (broken line). **E** Membrane potentials in response to the simultaneous presentation of the two stimulus modulations in **C** and **D**. The amplitude and phase modulations are aligned in time, representing situations where the animal exhibits JARs with decreasing frequency (left column, Df > 0) or increasing frequency (right column, Df < 0). Black lines in **E** indicate the actual membrane potential, gray lines indicate numeric sum of actual responses to component stimuli in **C** and **D**. This neuron strongly preferred the situation in the left column, giving deep excitatory synaptic potentials and action potentials (which have been removed for the sake of analysis and presentation in this figure). From Carlson and Kawasaki (2006)

ameters differs for the different senses of rotation but each parameter changes identically for both senses of rotation of the graph. These neurons do not interact with each other in the ELL, but project to common areas in the midbrain torus semicircularis (Kawasaki and Guo 1998). There, one finds 'sign-selective' neurons that are selective to the sense of rotation (Fig. 8A, B) (Kawasaki and Guo 2002). While some neurons prefer the clockwise rotation, other neurons prefer the counterclockwise rotation. Carlson and Kawasaki (2004, 2006) recorded from these 'sign-selective' neurons with extracellular and *in vivo* whole-cell techniques to reveal synaptic potentials that interact within these neurons for the differentiation of the time sequence of the modulations (Fig. 8C–E).

A singular presentation of an amplitude or phase modulation evoked relatively weak synaptic potentials in these midbrain neurons that were time locked to the modulation (Fig. 8C, D). The positive and negative peaks in synaptic potentials occurred at any time position depending on neurons indicating various latency or delay from the modulation peaks to the peaks in the synaptic potentials in these midbrain neurons. These delays are presumably due to various physiological mechanisms existing between the electroreceptor and these neurons. They include adaptation of transduction mechanisms, firing mechanisms of the cell membrane, and excitatory and inhibitory synapses. Due to the delays, the synaptic potential peaks in these midbrain neurons are aligned in time resulting in a maximum response to either a clockwise or a counterclockwise combination of input signals (Fig. 8C–E). Thus the time relations that are specific to a particular sense of

rotation (Fig. 7) are detected by coincidence of postsynaptic potentials in the midbrain neurons. Coincident arrivals of postsynaptic potentials are enhanced by voltage sensitive membrane processes to give stronger preference to the corresponding sense of rotation (Carlson and Kawasaki 2006).

4. Comparison of microsecond sensitive systems

Phase (time) difference signals on the order of microseconds are used for various behavioral functions in different types of weakly electric fishes. A gymnotiform fish, *Eigenmannia*, is a wave type electric fish that generates wave type EODs. *Eigenmannia* exhibits JARs that are very similar to those of *Gymnarchus*. Despite being independently evolved electric fishes, *Gymnarchus* and *Eigenmannia* use complex yet identical computational algorithms for their JARs. They both use the time pattern of amplitude and differential phase modulations to encode the difference in discharge frequencies (Heiligenberg and Bastian 1980; Heiligenberg 1991; Kawasaki 1996). Though the computational algorithms and the parallel and convergent processing of amplitude and phase information are shared, these fishes have neuronal mechanisms for differential phase computation in different brain areas. The circuit occurs in the hindbrain in *Gymnarchus*, but in the midbrain in *Eigenmannia.* Both fishes exhibit JARs with very small magnitudes of amplitude and differential phase modulation (Rose and Heiligenberg 1985; Kawasaki et al. 1988; Guo and Kawasaki 1997). The hyper accurate JARs, however, are small in regard to the magnitude of frequency shifts and slow in regard to time course. Since the function of the JARs is to increase the frequency of amplitude and time modulation by increasing the frequency difference between a fish's and its neighbor's EODs for jamming avoidance, such slow and small changes in EOD frequency would hardly function as jamming avoidance. The electrolocation of objects, however, demands such high sensitivities because the magnitude of the electric dipole field that the fish generates around itself attenuates as the inverse cube of the distance. The detection of resistive and capacitive objects at some distances requires high sensitivities. Thus, the extreme sensitivities to amplitude and time may have evolved both in *Gymnarchus* and *Eigenmannia* under the selection pressure for the detection of extremely small modulations of feedback signals for electrolocation objects away from the fish. This suggests that neuronal mechanisms for JARs evolved in both Mormyriformes and Gymnotiformes based on neuronal circuits preadapted to the electrolocation functions. The presence of a gymnotiform fish, *Sternopygus*, that does not perform JARs but possesses central neurons that respond to electrosensory stimuli that would evoke JAR in other species, supports this notion (Rose et al. 1987).

Gymnotiform pulse-type electric fishes perform a different type of jamming avoidance response. Rather than changing the discharge frequencies smoothly and gradually, pulse-type gymnotiform fish shift the firing times of a few EODs transiently to avoid synchronized firing of EODs with those from neighbors. Studies in *Brachyhypopomus* (Heiligenberg et al. 1978; Baker 1980) demonstrated that this jamming avoidance response does not require differential phase information. A temporal pattern of amplitude modulation alone, evoked in the sensory feed back signal, is sufficient. The midbrain of these fishes, however, possesses large neurons that are reminiscent of the differential phase comparison circuit in wave-

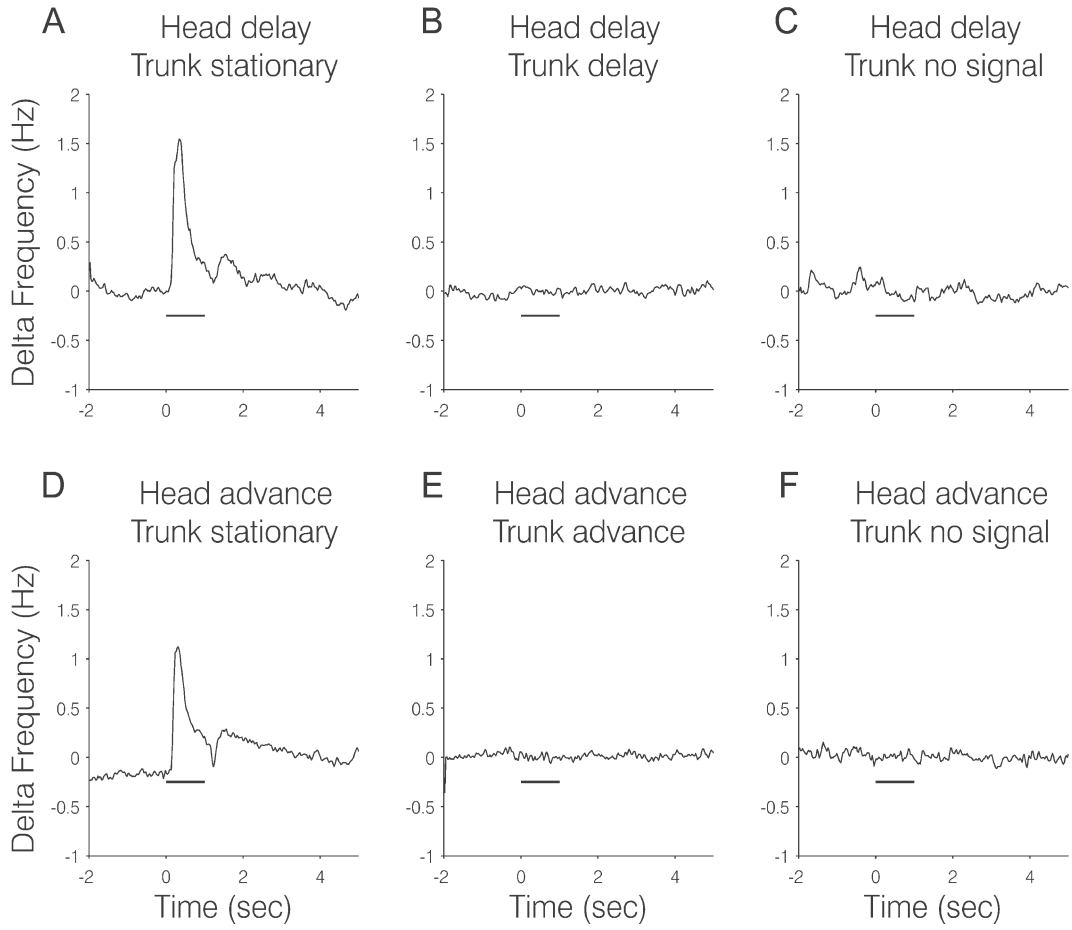


Fig. 9 Novelty responses to time differences in a pulse-type gymnotiform fish *Brachyhypopomus gauderio*. Curarized fish were placed in a phase chamber as in Fig. 6a. **A** 30 Hz pulse train was given simultaneously to both chambers up to time zero at which moment the pulse phase was either delayed or advanced by 200 µsec for one second (stimulus bars) (**A**, **B**, **D**, and **E**). In **C** and **F**, there was a stimulus only in the head chamber throughout the trace. Note that a transient frequency increasing behavior (novelty response) was observed only when there was a time difference between the head and trunk signals (**A** and **D**)

type gymnotiform fishes (Réthelyi and Szabo 1973). The actual use of the differential phase sensitivity in these fishes has not been known so far. Our preliminary experiments, however, indicate that *Brachyhypopomus* shows a behavioral response to phase differences (Fig. 9). *Brachyhypopomus* was placed in the phase chamber as in Fig. 6A and given a continuous pulse train at ~30 Hz. When a time difference of approximately 200 µsec was introduced between pulses stimulating different parts of the body, the fish transiently raised the EOD frequency. This type of frequency raising behavior is known as the 'novelty' response that occurs in response to changes in electrolocation signals. The threshold value for time difference for a novelty response is less than 10 µsec (Kawasaki and Matsushita 2009).

Mormyrid pulse type electric fishes also

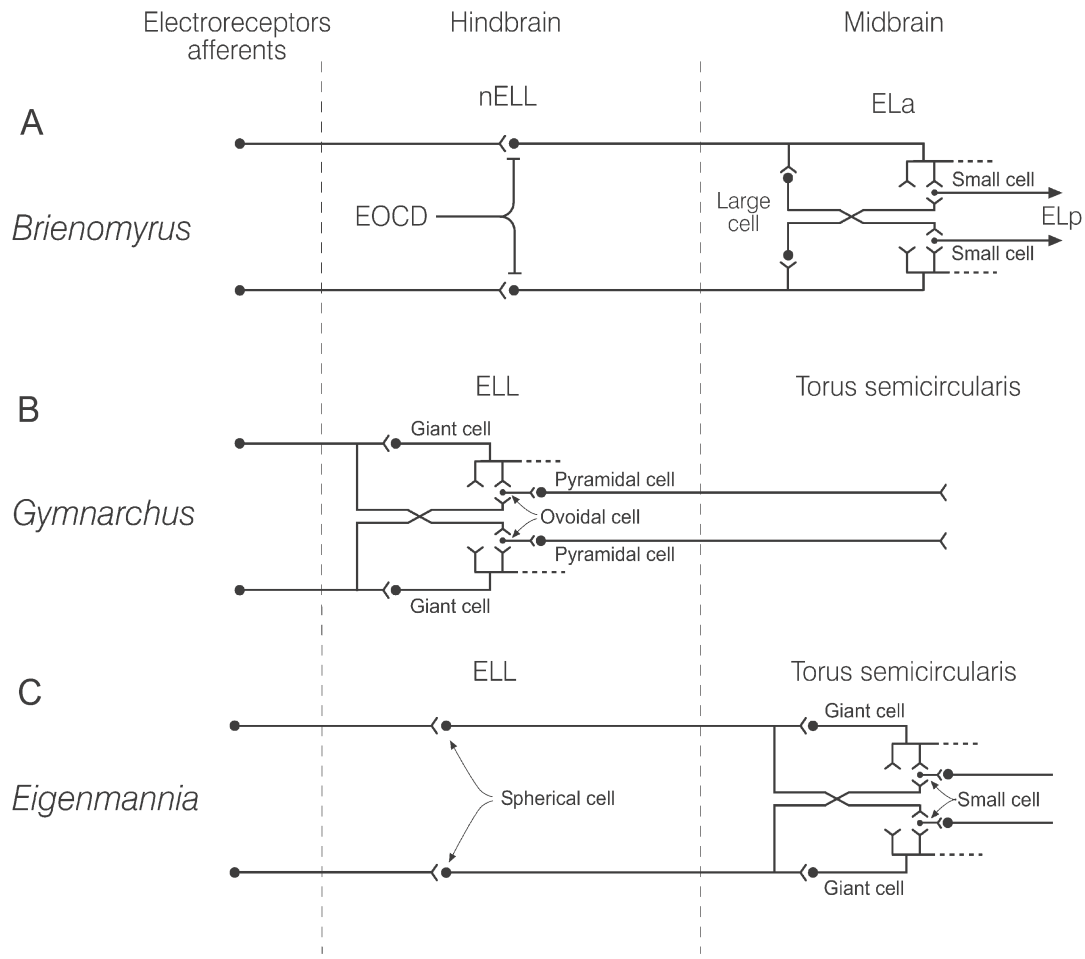


Fig. 10 Comparison of phase (time) comparison circuits in a mormyriform pulse-type fish *Brienomyrus* (**A**), mormyriform wave-type fish, *Gymnarchus* (**B**), and a gymnotiform wave-type fish *Eigenmannia* (**C**). The two parallel channels in each row represent signal paths from different areas of the body between which time difference is detected. The small cell in *Brienomyrus* and *Eigenmannia*, and the ovoidal cell in *Gymnarchus* are thought to be the phase (time) comparators which compare the phases (times) of firing of input neurons. Note the similar topographical organization of the time comparison circuits in all species

possess a time (phase) comparison system. The system does not seem to be involved in electrolocation or in the jamming avoidance response, but performs species and sex recognition by waveform analyses of short EOD pulses (Hopkins 1986). A type of electroreceptor, the Knollenorgan, fires one action potential in response to one phase of an EOD pulse (up or down stroke of a pulse). Knollenorgans at different locations on the body surface may be stimulated by different phases of an EOD pulse. Thus, the pulse duration is encoded by the time differences between different afferent fibers from the Knollenorgans. The duration of the EOD pulses in mormyrid electric fishes ranges from 100 µsec to 30 msec. While action potentials generated by self-stimulation due to the fish's own EODs are blocked in the nucleus of the electrosensory lateral line lobe (Zipser and Bennett 1976; Bell et al. 1983), those evoked by the neighbor's EODs reach

the midbrain where the time differences are detected. The small cells in the nucleus exterolateralis pars anterior (ELa) are sensitive to time differences between signals at different parts of the body (Friedman and Hopkins 1995, 1998; Amagai et al. 1998; Xu-Friedman and Hopkins 1999).

Fig. 10 summarizes the phase (time) coding systems of some of above-mentioned fishes which exhibit striking similarities despite their distant phylogenetic relationships. The shared features include (1) large-diameter and fast-conducting axons in the time pathways; (2) adendritic somata in phase-locked neurons; and (3) existence of mixed synapses (electrical and chemical) both at phase-conserving synapses that conserve the spike time sequence and at time-comparing synapses that drive time comparator postsynaptic neurons. Fast conduction is believed to contribute to accurate conduction of time-locked firing with minimal jitter. Adentritic postsynaptic cells are probably an adaptation for fast electrotonic propagation and integration of synaptic potentials (Carr 1993; Carr and Friedman 1999). Another striking similarity is that time-locked axons bifurcate into direct and indirect paths to provide inputs to time comparators, and the time comparator neurons (small cells in *Eigenmannia* and *Brienomyrus* and ovoidal cells in *Gymnarchus*) receive inputs from the indirect and direct paths for time comparison.

Comparative analyses of the electrosensory systems reveal signs of evolutionary development and alteration of the functions of nervous systems. The frequency analyses in the JAR rely on amplitude and timing sensing systems that had evolved and currently still functions for the electrolocation behavior. The neuronal mechanism for the JARs in electric fishes is an example where a new system is piled on top of old ones in the phylogenetic evolution. Such a design feature may be found in both biological and human engineered systems (Marcus 2008).

References

Amagai S, Friedman MA, Hopkins CD (1998) Time coding in the midbrain of mormyrid electric fish. I. Physiology and anatomy of cells in the nucleus exterolateralis pars anterior. J Comp Physiol A 182(2): 115–130

Baker CLJ (1980) Jamming avoidance behavior in gymnotoid electric fish with pulse-type discharges: Sensory encoding for a temporal pattern discrimination. J Comp Physiol A 136: 165–181

Bell CC, Libouban S, Szabo T (1983) Pathways of the electric organ discharge command and its corollary discharges in mormyrid fish. J Comp Neurol 216: 327–338

Bullock TH (1970) The reliability of neurons. J Gen Physiol 55: 565–584

Bullock TH, Behrend K, Heiligenberg W (1975) Comparison of the jamming avoidance responses in Gymnotoid and Gymnarchid electric fish: A case of convergent evolution of behavior and its sensory basis. J Comp Physiol 103: 97–121

Carlson BA, Kawasaki M (2004) Nonlinear response properties of combination-sensitive electrosensory neurons in the midbrain of *Gymnarchus niloticus*. J Neurosci 24(37): 8039–8048

Carlson BA, Kawasaki M (2006) Stimulus selectivity is enhanced by voltage-dependent conductances in combination-sensitive neurons. J Neurophysiol 96: 3362–3377

Carr CE (1993) Processing of temporal information in the brain. Annu Rev Neurosci 16: 223–243

Carr CE, Friedman MA (1999) Evolution of time coding systems. Neural Comput 11(1): 1–20

Friedman MA, Hopkins CD (1995) Evidence for mechanisms of temporal analysis in the knollenorgan electrosensory system of mormyrid fish. In: M Burrows et al. (eds) Nervous systems and behavior. Stuttgart: Georg Thieme Verlag, pp. 419

Friedman MA, Hopkins CD (1998) Neural substrates for species recognition in the time-coding electrosensory pathway of mormyrid electric fish. J Neurosci 18(3): 1171–1185

Guo Y-X, Kawasaki M (1997) Representation of accurate temporal information in the electrosensory system of the African electric fish*, Gymnarchus niloticus*. J Neurosci 17(5): 1761–1768

Heiligenberg W (1975) Electrolocation and jamming avoidance in the electric fish *Gymnarchus niloticus* (Gymnarchidae, Mormyriformes). J Comp Physiol 103: 55–67

Heiligenberg W (1991) Neural nets in electric fish. Cambridge, MA: The MIT Press

Heiligenberg W, Baker C, Bastian J (1978) The jamming avoidance response in gymnotoid pulse-species: A mechanisms to minimize the probability of pulse-train coincidence. J Comp Physiol 124: 211–224

Heiligenberg W, Bastian J (1980) The control of *Eigenmannia*'s pacemaker by distributed evaluation of electroreceptive afferences. J Comp Physiol A 136: 113–133

Hopkins CD (1986) Behavior of Mormyridae. In: TH Bullock and W Heiligenberg (eds), Electroreception. New York: John Wiley & Sons, pp. 527–576

Kawasaki M (1993) Independently evolved jamming avoidance responses employ identical computational algorithms: a behavioral study of the African electric fish, *Gymnarchus niloticus*. J Comp Physiol A 173(1): 9–22

Kawasaki M (1996) Comparative analysis of the jamming avoidance response in African and South American wave-type electric fishes. Biol Bull 191(1): 103–108

Kawasaki M, Guo Y-X (1996) Neuronal circuitry for comparison of timing in the electrosensory lateral line lobe of the African wave-type electric fish *Gymnarchus niloticus*. J Neurosci 16: 380–391

Kawasaki M, Guo Y-X (1998) Parallel projection of amplitude and phase information from the hindbrain to the midbrain of the African electric fish *Gymnarchus niloticus*. J Neurosci 18(18): 7599–7611

Kawasaki M, Guo Y-X (2002) Emergence of temporal-pattern sensitive neurons in the midbrain of weakly electric fish *Gymnarchus niloticus*. J Physiol (Paris) 96: 531–537

Kawasaki M, Matsushita A (2009) Behavioral sensitivity to time differences in the electrosensory system of a pulse type gymnotiform fish. Neuroscience Abstracts, Chicago: Society for Neuroscience Online

Kawasaki M, Rose GJ, Heiligenberg W (1988) Temporal hyperacuity in single neurons of electric fish. Nature 336: 173–176

Marcus GF (2008) Kluge: the haphazard construction of the human mind. Boston: Houghton Mifflin.

Matsushita A, Kawasaki M (2004) Unitary giant synapses embracing a single neuron at the convergent site of time-coding pathways of an electric fish, *Gymnarchus niloticus*. J Comp Neurol 472(2): 140–155

Matsushita A, Kawasaki M (2005) Neuronal sensitivity to microsecond time disparities in the electrosensory system of *Gymnarchus niloticus*. J Neurosci 25: 11 424–11 432

Miles RN, Hoy RR (2006) The development of a biologically-inspired directional microphone for hearing aids. Audiol Neurootol 11(2): 86–94

Réthelyi M, Szabo T (1973) A particular nucleus in the mesencephalon of weakly electric fish, *Gymnotus carapo* (Gymnotidae, Pisces). Exp Brain Res 17: 229–241

Rose GJ, Heiligenberg W (1985) Temporal hyperacuity in the electric sense of fish. Nature 318: 178–180

Rose GJ, Keller CH, Heiligenberg W (1987) "Ancestral" neural mechanisms of electrolocation suggest a substrate for the evolution of the jamming avoidance response. J Comp Physiol A 160: 491–500

von der Emde G (1992) Electrolocation of capacitive objects in four species of pulse-type weakly electric fish. II. Electric signaling behaviour. Ethology 92: 177–192

Xu-Friedman MA, Hopkins CD (1999) Central mechanisms of temporal analysis in the knollenorgan pathway of mormyrid electric fish. J Exp Biol 202(10): 1311–1318

Zipser B, Bennett MVL (1976) Interaction of electrosensory and electromotor signals in lateral line lobe of a mormyrid fish. J Neurophysiol 39: 713–721

Bioinspired sensors, sensor materials and fabrication

VI

Synthetic materials for bio-inspired flow-responsive structures

23

Michael E. McConney, Vladimir V. Tsukruk

Contents

Abstract

Fish rely on flow sensors for predator avoidance, hunting, rheotaxis, and other vital functions. The cupula on top of the sensory cells of these sensors is a structure that couples the water flow close to the fish and the receptor cells by efficiently absorbing the mechanical energy contained in the stimulus. The structure and mechanical properties of the fish cupula were investigated in order to develop a bio-inspired analogue to be used for synthetic flow receptors. After having studied the biological original structure, an engineered analogue of the cupula was developed. The addition of the artificial cupula to the flow sensors improved both the response magnitude (sensitivity) and the detection threshold by almost 40 times. In the end, the bio-inspired modification resulted in engineered sensors with capabilities comparable to those of the fish flow sensors.

Vladimir V. Tsukruk
Georgia Institute of Technology
School of Materials Science and Engineering
771 Ferst Drive, Atlanta, GA 30332-0245, USA
e-mail: vladimir@mse.gatech.edu

1. Introduction

Bio-inspired design of responsive materials and structures involves an understanding of the design principles and strategies employed by biological systems, in order to benefit the development of engineered systems with enhanced properties. From a materials perspective, biology offers an almost limitless source of novel approaches capable of arousing innovation in various aspects which include fabrication, design, and functionality (see various relevant chapters in this volume).

Biologically inspired design can catalyze development through the use of a bi-directional approach to problem solving, where both the problem and potential solutions are analyzed simultaneously (Bar-Cohen 2006). Although there are many examples of success stories involving bio-inspired design, the rules and the methods of this approach are still evolving. In order to fully leverage this bi-directional approach it is important to understand the fundamental

principles. Generally, nature's approach is to continuously improve past designs and make changes to solve new challenges; whereas engineers, often literally start their designs from the drawing board (Bar-Cohen 2006).

Flow measurement is a basic area of measurement, utilizing many different physical relationships (Nguyen 1997, Darby 2001). Although simply measuring flow velocity is a straightforward process, 3D flow visualization with high spatial resolution (velocity, pressure, and vorticity fields) is not nearly as simple. From a practical viewpoint, there is a large demand for a compact flow-visualization system that is capable of acting as a detection system for a variety of applications, including the guidance of autonomous underwater vehicles. A major advantage of a passive-array technique compared to more common bulk-based (free-flow) techniques, such as Doppler shift, is the ability for dramatic miniaturization, low energy consumption, and the absence of any energy emission into the environment. A major disadvantage of flow sensors like those discussed in this chapter is that the sensing is taking place near the stagnant boundary layer of the surface, which significantly increases the need for high sensitivity and relatively tall sensors that still do not significantly interfere with the flowing environment. Furthermore, relatively large arrays of individual sensors are required to accurately visualize the surrounding flow. This requirement leads to signal processing issues, where increasing demands on computing limit the portability of such systems.

Work presented in this chapter includes a brief overview of our recent studies of fish water flow receptors and the subsequent development of bio-inspired synthetic structures to improve flow sensor performance. We summarize a series of recent investigations from our group which have been published in a number of recent reports (Peleshanko et al. 2007, Anderson et al 2008, McConney et al. 2009a,b).

2. Developing underwater flow sensing materials

The artificial hair-like flow sensors developed in Chang Liu's group at Northwestern University are piezoresistive sensors inspired by fish lateral line flow receptors (Liu 2007). The current microfabricated hair sensors consist of a microscopic polymer "hair" with a diameter below 100 micrometers and photolithographically-mounted onto a silicon-based piezoresistive microcantilever. The whole sensor system is covered by parylene, which acts as a robust water proof coating. The flow-derived signal is generated by the deflection of the hair under flow and the corresponding resistance change of the piezoresistive microcantilever due to its bending. In their original design these sensors are robust and have a minimum detection threshold of flow velocities above 0.2 mm/s. This threshold is well below current commercial devices, but still much higher than that required for the visualization of fine flow structures with microscopic resolution. Minimizing the detection threshold is a key factor for many practical reasons, including the enhancement of the spatial range of underwater monitoring systems and the precise determination of the underwater environment.

As is known from numerous biological studies, fish rely on flow receptors for several important tasks, including navigating, hunting prey, rheotaxis and schooling. Furthermore, the flow receptors allow blind fish to avoid colliding with stationary objects (Hassan 1989). Fish have developed

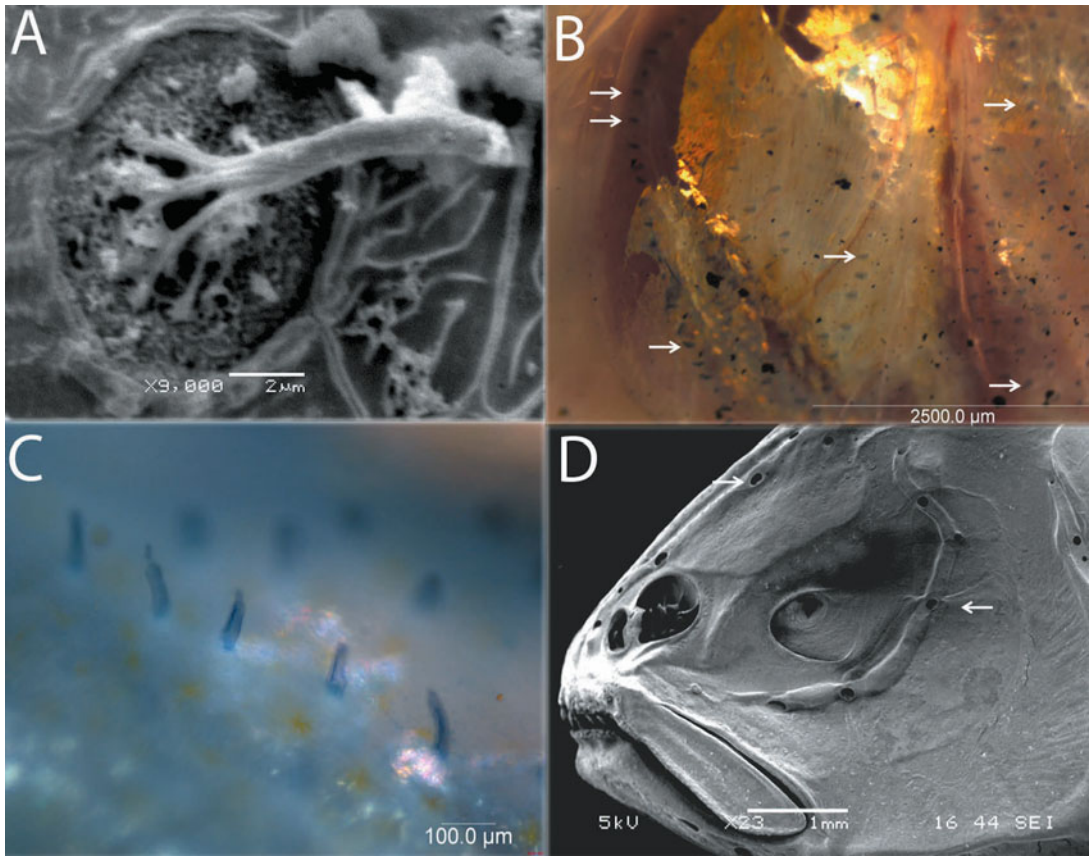


Fig. 1 **A** SEM image of a fish neuromast with the cupula removed by enzyme treatment to visualize the stereovilli and kinocilium; **B** optical micrograph of fish body surface with stained superficial cupulae (arrows indicate a few examples); **C** optical micrograph of array of stained cupulae at high magnification; **D** SEM image of fish head showing the canals of the lateral line (arrows indicate examples of canal openings)

the ability to sense flow rates in water as low as several tens of a micrometer per second (Kroese et al. 1978, Coombs and Janssen 1989). The lateral line is a system of flow transducing neuromasts. These are either located on the body surface and referred to as free neuromasts, or arranged inside canals with series of pores opening to the outside and then called canal neuromasts (Fig. 1). A neuromast, the basic flow sensing unit in fish, typically includes from 20 to 1000 mechanosensitive hair cells (Coombs and van Netten 2006). These are covered by a very compliant bio-hydrogel structure called a cupula (Bleckmann et al. 2001, McHenry and van Netten 2007; see also chapter I,2 by H Bleckmnn et al. and chapter VII,9 by N Izadi and G Krijnen). Cupulae are usually around 100–1 000 µm high, but their size and mechanical properties have been shown to vary greatly in different species (Teyke 1990). They efficiently couple the hair cells (sensory cells) to the surrounding water flow by increasing the drag of the neuromasts, thereby enhancing stimulus absorption and transmission. Our work was aimed at understanding the design that endows fish with the ability to sense flow with a very low minimum detection threshold. Following preliminary studies and the examination

of the literature it became apparent that the cupula was a key functional structure of fish flow receptors. Our studies therefore focused on understanding and implementing the role of the cupula in synthetic sensors.

The dimensions and aspect ratio of superficial cupulae of blind cave fish were measured using atomic force microscopy (AFM), scanning electron microscopy (SEM), fluorescence optical microscopy (FOM), and conventional optical microscopy. As can be seen from FOM images the superficial cupulae (neuromasts) are either randomly arranged over large areas of the fish body surface or arranged in linear arrays with regular spacing between the neuromasts of 100–200 micrometers (Fig. 1). Superficial cupulae stained with fluorescent dye possess a flag-like shape with an average height of 100 µm, a width of 26 µm, and an aspect ratio of about 4.0.

In order to guide the development of a synthetic cupula, the mechanical properties of superficial cupulae in blind cave fish have also been characterized. They were directly measured using fluid-based surface force spectroscopy with a microscopic colloidal probe directly on the cupula using typical AFM probing procedures (Chizhik et al. 1998). A SEM image of the surface region with a cupula covered with a mucus layer is presented in Fig. 2. The fish scales carrying cupulae were washed to remove the mucus layer prior to AFM probing. The elastic modulus of such cupulae was measured in water using a microcantilever with attached silica beads 5 to 10 µm in diameter. The cantilever was operated in force-volume mode to collect 256 force-distance curves from a selected surface area (Fig. 2). A representative plot with the coordinates showing penetration depth versus the applied load for normal loads below 1 nN and an overall deformation below 100 nm is presented in Fig. 2. The loading curves are highly non-linear due to the viscous contribution of highly viscoelastic hydrogel materials associated with the time-dependent mechanical response.

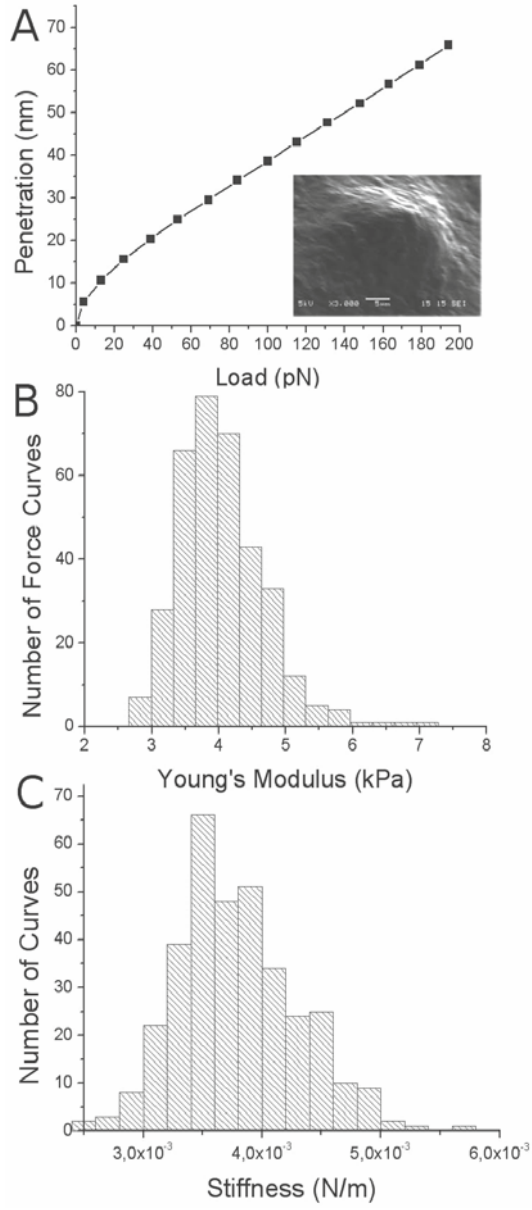


Fig. 2 **A** AFM loading data for biological cupula at very low loads applied perpendicular to the surface of the fish, inset is an SEM micrograph of a dried fish cupula **B**, **C** Histograms of surface distribution of elastic modulus and surface stiffness for cupula surface obtained from AFM micromapping

The two-element Voight viscoelastic model was combined with the Hertzian con-

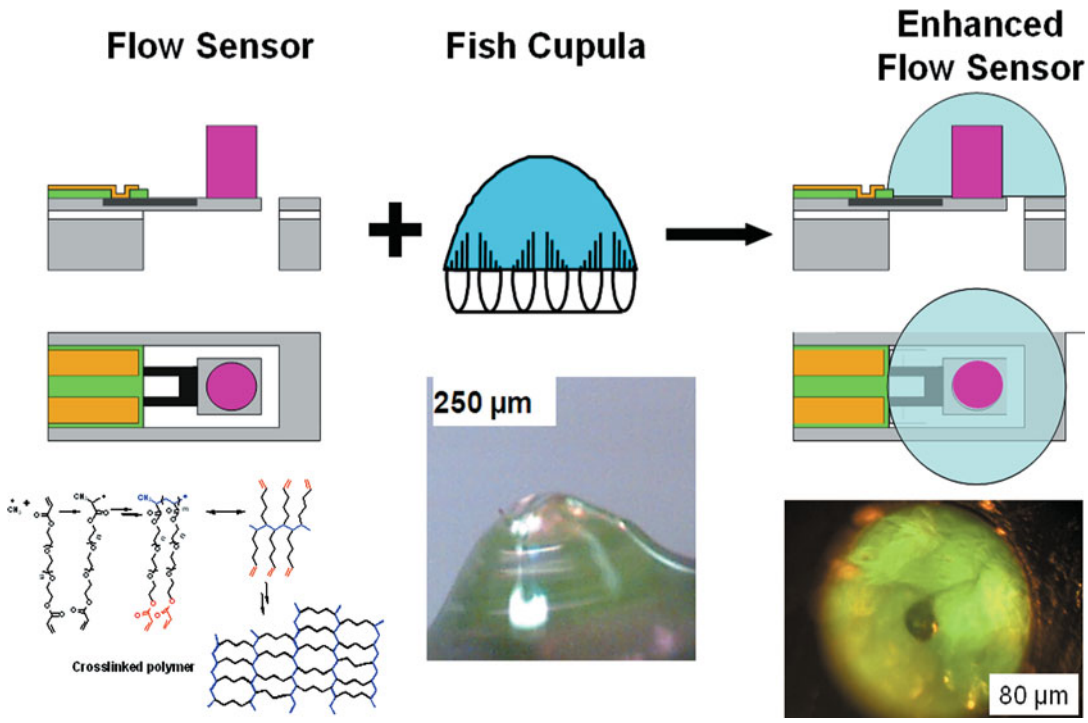


Fig. 3 **Left** Fabrication procedure and chemical structure of synthetic PEG cupula grown directly on the artificial hair sensor. Both the side and top-view schematics are presented for the whole sensory system. The component shown in yellow represents the conductive contact that completes the conductive circuit by making contact with the piezoresistive doped silicon shown in black. The components shown in green and grey are passive layers. The signal absorbing hair made from SU-8 photoresist is drawn in pink. **Center** and **right** Side view and top-view optical images of stained dome-shaped synthetic cupulae with imbedded hair

tact model to fit the non-linear loading data for the cupula. The analysis of the data array has generated histograms of the surface distribution for the elastic modulus and the surface stiffness for selected surface areas of several micrometers across (Fig. 2). These data show relatively narrow distributions indicating a homogeneous composition of the cupula biohydrogel material. The average surface stiffness of about 4×10^{-3} N/m is typical for very compliant biohydrogel material highly swollen in an aqueous environment. The biohydrogel material has an elastic modulus from 4 to 10 kPa and a relaxation time of 0.4–0.5 s, which are values characteristic of swollen hydrogels.

The synthetic cupulae with mechanical properties similar to those of the fish cupulae were fabricated by photo-crosslinking tetra-acrylate functionalized PEG (polyethylene glycol) directly on top of the microfabricated sensor (Fig. 3). PEG was deposited directly onto the hair sensor and then exposed to ultraviolet light (wavelength 365 nm) of varying intensity and exposure time. Photomasks with proper dimensions were used to confine the crosslinking of the PEG to certain areas through selective light exposure. The resulting synthetic cupula has a dome-like shape with the hair completely covered by the PEG hydrogel. It is important to note that when the PEG solution is dropcast onto the hair structure, the surface at the base of the hair is wet, resulting in a

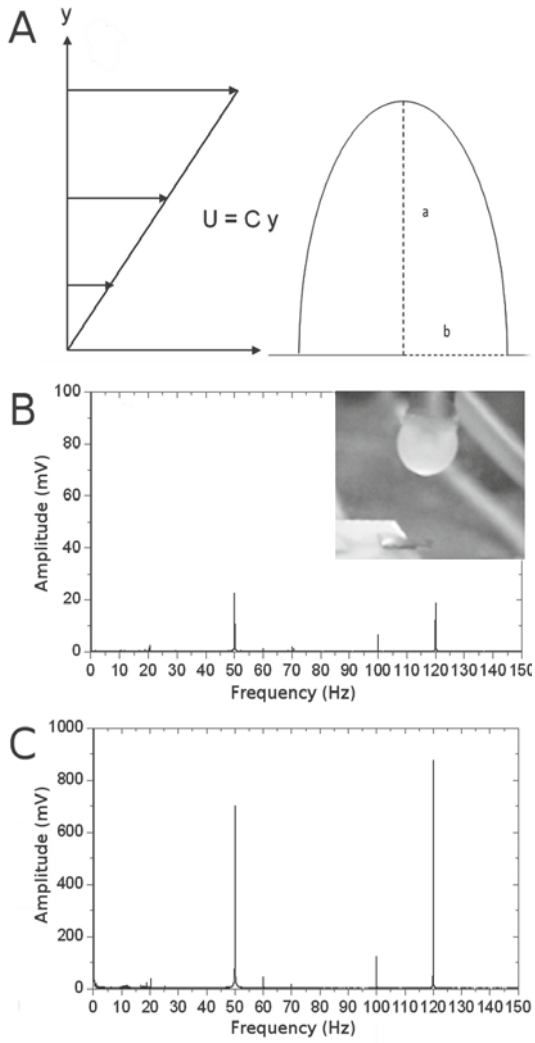


Fig. 4 **A** Distribution of the flow velocity, U, as a function of distance, y, from a surface and a schematic of an elliptical cross-section of a synthetic cupula (a and b, long and short axes), used to gain theoretical estimates of the sensitivity difference between a bare and cupula covered hair sensor due only to the change in cross-sectional area (plot provided by JAC Humphrey). **B**, **C** Fourier transforms of experiments using a dipole oscillating at 50 Hz as a source of water flow for a bare hair sensor (**B**) and one capped by a cupula (**C**) (other peaks in the graphs represent contributions from laboratory and equipment characteristic frequencies, except for harmonics at 100 Hz). **B, inset** Arrangement of oscillating dipole (metal ball in the center of the image) in the water tank with hair sensory system fixed by two holders below ball

dome-like structure (Fig. 3). The base diameter of these dome-shaped synthetic cupulae was within 1 to 2 mm and their aspect ratio did not exceed 1.

The flow sensor testing was carried out under water in a special tank by shaking a dipole placed at set distances from the sensor surface (Fig. 4). The dipole amplitude and frequency were controlled and monitored, while simultaneously monitoring the Fourier transform of the sensor's piezoresistive output. The dipole amplitude was varied to obtain the sensor's sensitivity to flow rate. The minimum threshold stimulus has been measured by lowering the dipole excursion amplitude until the sensor output could not be detected anymore. Following the measurements with the initial sensor with a bare hair, the sensor was encapsulated in a dome-shaped cupula as described above and tested under identical flow conditions. This experimental routine was implemented to avoid large variations of initial sensitivity between sensors and to ascertain measurements under identical experimental conditions. Figure 4 demonstrates the results of the Fourier transformation from two such experiments at 50 Hz dipole oscillation for the initial and the encapsulated hair sensors. As becomes apparent from the direct comparison, the amplitude of the main oscillation signal at 50 Hz dramatically increased, by a factor of 25 on average, after the addition of the synthetic hydrogel cupula around the hair sensor.

Correspondingly, as we observed in these measurements at the lowest possible amplitude of oscillation, the unaltered sensors tested here had minimum flow velocity detection thresholds of around 0.2 mm/s in accordance with previous results from Liu's group (Liu 2007). However, after the synthetic hydrogel cupula was fabricated onto the sensors, the threshold velocities decreased by a factor of ca. 1/3 to roughly 0.075 mm/s. Theoretical estimates derived from the nor-

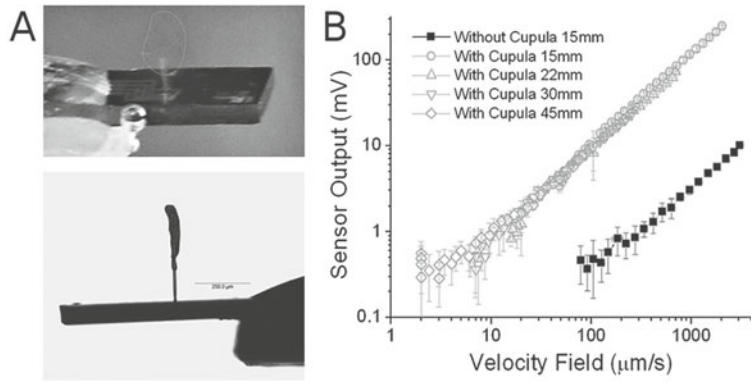


Fig. 5 **A** Tall superficial synthetic cupula on hair sensor in dry state (bottom) and in water (top). **B** Sensor output vs flow velocity for bare and cupula-capped hair sensors. The distance of oscillating dipole from the sensor was varied between 15 mm and 45 mm (see text)

mal distribution of the velocity field and the elliptical cross-section of dome-shaped cupulae indicated that the expected signal amplification due to the increased cross-sectional area of the hair-cupula sensor accounted for only about half of the actual signal amplification (Fig. 4B). The discrepancy can be attributed to an enhanced signal absorption due to an additional frictional factor specifically associated with the high permeability of the hydrogel network and intermolecular interactions of the hydrogel with the flowing water.

The focus of the next step to be taken was to improve the previous dome-like cupula's performance by fabricating a higher-aspect ratio cupula, much like that of flag-like fish superficial cupulae (McHenry and van Netten 2007). We used an aspect ratio of 4.0 (Fig. 1) as a general guide for the development of high-aspect ratio synthetic cupulae (Fig. 5). In order to fabricate tall synthetic hydrogel cupulae with a flag-like shape we developed a controlled drop-casting method (McConney et al. 2009a). To facilitate this endeavor a 3-axis micro-positioner was used to precisely position a microsyringe filled with the functional PEG macromonomer solution directly above the hair of the sensor. Then several drops of PEG solution were precisely dropped onto the hair without wetting the base surface. After the deposition of precursor solution, the fast evaporation of solvent resulted in fast drying and partial polymerization. These fast processes fixed the initial shape of the dried hydrogel material and made its robust grafting to the hair sensor possible.

The resulting tall and stable synthetic cupula preserved its shape in dry state and did not collapse after having been kept for several weeks in a modestly humid lab environment. The tall cupulae can be significantly swelled by placing them into water thereby significantly increasing their cross-sectional area and their interaction with the flow environment (Fig. 5). The fabrication method applied provided a certain degree of control of the height and width of the cupula by controlling the number of drops and the volume of each drop, respectively. Micromechanical testing indicated that this synthetic cupula material was softer (E ~10's Pa) than that of the photopolymerized dome-shaped synthetic cupulae due to the low density of the regular covalent crosslinking network.

The performance of the higher-aspect ratio synthetic cupula as a sensing enhancement structure was tested in a similar manner as that of the dome-shaped cupula, using the standard dipole test (Fig. 5). There was an impressive difference in sensor performance before and after the addition of the bio-inspired structure: the sensitivity improvement was more than 60 times for the high-aspect ratio cupula, which is

a tremendous improvement, much higher than that observed for the dome-shaped cupulae discussed above (Fig. 5). In fact the sensitivity was so high that the distance between the oscillating dipole and the sensor had to be increased from 15 mm to 45 mm in order to quantify the minimum detectable threshold because the shaker was incapable of small enough amplitudes to perform all measurements at 15 mm away (Fig. 5). The high-aspect ratio cupula improved the minimum detectable velocity, which was about 100 µm/s for the initial bare hair sensor to below 3 µm/s for the sensor with the flag-like cupula. The dome-shaped cupula resulted in an intermediate minimum detectable threshold velocity of 75 µm/s, which was significantly outperformed by the high-aspect cupula by a factor of 1/25 on average.

The unprecedented minimum detectable flow is remarkable considering that the limit of detectable flow of the initial bare hair sensors was 0.1–0.2 mm/s. It is even lower than the minimum detectable flows of 20–40 µm/s that were measured in different fish and suggested to represent a uniquely high sensitivity (Kroese et al. 1978). Therefore, by the application of bio-inspired sensor super-structures, we were able to truly rival the performance of our biological models.

Conclusions

This chapter presents the authors' efforts in the area of bio-inspired approaches regarding the design of structured soft materials for sensing applications. Specifically, we focused on fish-inspired synthetic structures made of functionalized polymeric hydrogel materials for enhancing underwater flow sensing. By studying fish and understanding that the hydrogel-like cupula of its neuromasts was a specialized structure that increased the sensitivity of hair-cells for water flow stimuli, we were able to demonstrate the value of a bio-inspired approach for the development of an engineered cupula applied to a synthetic hair-like flow sensor. With physical properties close to those of the biological cupulae, the artificial cupulae led to dramatically enhanced flow sensing capabilities even exceeding those observed for selected fish examples. The synergies resulting from an interaction between biology and engineering are clearly demonstrated by this study.

Acknowledgements

The authors would like to thank Prof. S. Coombs for her collaboration on the work with fish, the late Prof. J. A. C. Humphrey for collaboration on the calculations related to the model shown in Fig. 4A and Prof. C. Liu for the supply of artificial hair sensors. We would also like to sincerely thank Prof. Barth for his helpful comments and suggestions with the manuscript. We would also like to thank C. Schaber, as well as former and current members of Tsukruk's research group, Dr. S. Peleshanko, Dr. M. Ornatska, Dr. M. Julian, K. Anderson, and Dr. M. Lemieux for their invaluable contributions, discussions, and assistance. This work was supported by DARPA (BioSenSE), AFOSR, and AFRL.

References

Anderson KD, Lu D, McConney M, Han T, Reneker, DH, Tsukruk VV (2008) Hydrogel microstructures combined with electrospun fibers and photopatterning for shape and modulus control. Polymer 49: 5284–5293

Bar-Cohen Y (2006) Introduction to biomimetics: The wealth of inventions in nature as an inspiration for human innovation. In: Bar-Cohen Y (eds) Biomimetics: biologically inspired technologies. CRC Press, Boca Raton, pp 2–40

Bleckmann H, Mogdans J, Dehnhardt G (2001) Lateral line research: the importance of using natural stimuli in studies of sensory systems. In: Barth FG, Schmid A (eds.) Ecology of sensing. Springer, Berlin Heidelberg New York, pp. 149–168

Chizhik SA, Huang Z, Gorbunov VV, Myshkin NK, Tsukruk VV (1998) Micromechanical properties of elastic polymeric materials as probed by scanning force microscopy. Langmuir 14: 2606–2609

Coombs S, Janssen J (1989) Peripheral processing by the lateral line system of the mottled sculpin

(*Cottus bairdi*). In: Coombs S, Görner P, Münz H (eds.) The mechanosensory lateral line: neurobiology and evolution. Springer-Verlag, New York, pp 299–319

Coombs S, van Netten SM (2006) The hydrodynamics and structural mechanics of the lateral line system. Fish Physiol 23: 103–140

Darby R (2001) Chemical engineering fluid dynamics. 2nd ed. Marcel Dekker, New York

Hassan ES (1989) Hydrodynamic imaging of the surroundings by the lateral line of the blind cave fish (*Anoptichthys jordani*) In: Coombs S, Görner P, Münz H (eds.) The mechanosensory lateral line: neurobiology and evolution. Springer-Verlag, New York, pp. 218–227

Kroese ABA,Van der Zalm JM, Van der Berken J (1978) Frequency response of the lateral-line organ of *Xenopus laevis*. Pflüg Arch Eur J Physiol 375: 167–175

Liu C (2007) Recent developments in polymer MEMS. Adv Mater 19: 3783–3790

McConney ME, Chen N, Lu D, Hu HA, Coombs S, Liu C, Tsukruk VV (2009a) Biologically inspired design of hydrogel-capped hair sensors for enhanced underwater flow detection. Soft Matter 5: 292–295

McConney M, Anderson KD, Brott LL, Naik RR, Tsukruk VV (2009b) Bioinspired material approaches to sensing. Adv Funct Mater 19: 2527–2544

McHenry MJ, van Netten SM (2007) The flexural stiffness of superficial neuromasts in zebrafish (*Danio rerio*) lateral line. J Exp Biol 210: 4244–4253

Nguyen NT (1997) Micromachined flow sensors – a review. Flow Meas Instrum 8: 7–16

Peleshanko S, Julian MD, Ornatska M, McConney ME, LeMieux MC, Chen N, Tucker C, Yang Y, Liu C, Humphrey JAC, Tsukruk VV (2007) Hydrogel-encapsulated microfabricated haircells mimicking fish cupula neuromast. Adv Mater 19: 2903–2909

Teyke T (1990) Morphological differences in neuromasts of the blind cave fish *Astyanax hubbsi* and the sighted river fish *Astyanax mexicanus*. Brain Behav Evol 35: 23–30

Polyelectrolyte hydrogels as electromechanical transducers

24

Katsiaryna Prudnikova, Marcel Utz

Contents

Abstract

Polyelectrolyte hydrogels consist of cross linked, charged polymer chains swollen in water or another highly dielectric solvent. They represent a salt solution in which ions of one type (usually, anions) are bound to the polymer backbone, while the other ions (usually, cations) are freely mobile in the solvent. This difference in mobility brings about interesting electromechanical properties. In this contribution, we discuss our recent research concerning the application of polyelectrolyte hydrogels as sensor materials.

A new indentation method for the quantitation of the electromechanical effect in polyelectrolyte hydrogels has been developed, and used for a systematic study of the relationship between the spatial charge density and the electromechanical coupling constant. Prototypes of microfluidic devices applying such gels as pressure transducers have been built, and preliminary experimental data of their performance is presented.

Marcel Utz
University of Virginia
Center for Microsystems for the Life Sciences
Department of Mechanical Engineering
104 Engineer's Way, Charlottesville, VA 22904
e-mail: mu3q@virginia.edu

1. Introduction

Polyelectrolyte hydrogels are linear, slightly cross-linked polymers containing ionizable groups on the backbone. When swollen in water or another solvent with high dielectric constant, the ionizable groups dissociate, producing a salt solution in which ions of one kind (usually negative) are attached to a polymer network, while the others (usually positive) are freely mobile in the solvent. This drastic difference in mobility leads to the electromechanical coupling effects that are the topic of this contribution.

Charged polymer systems are ubiquitous in nature, and are finding increasingly wide-

spread applications in technology. Tanaka et al. (1982) have shown that charged, cross-linked polymers swollen in water undergo large changes in volume and shape upon exposure to an electric field. This has inspired a large body of research, both experimental and theoretical, focused on studying the deformation of polyelectrolyte gels in response to various stimuli, such as changes in pH, temperature, and salinity (Beltran et al. 1991, Doi et al. 1992, de Gennes et al. 2000, De et al. 2002, Yao and Krause 2003, Wallmersperger et al. 2004, Yamaue et al. 2005). pH- and temperature-responsive hydrogels have been used as "intelligent" valves in microfluidic systems (Beebe et al. 2000, Richter et al 2003, Bassetti 2005). Electromechanical effects in polyelectrolytes bear interesting similarities to processes in biological systems. For example, the ability of perfluorinated ionomers such as Nafion to convert electrical inputs into a mechanical output has been likened to the stimulated contraction of muscle fibers (Shahinpoor et al.1998; Nemat-Nasser 2002; Shahinpoor and Kim 2002). Due to their potential application in robotics, these systems have been investigated extensively over the last decade. They differ from the gels under study in this contribution in that their water content is much lower (on the order of a few percent). Even though it seems likely that their operation is based on similar physical principles as the much more highly swollen hydrogels, but at the present time this is not yet entirely clear.

On a physical level, the electromechanical coupling in hydrogels is fully reversible, in the sense that mechanical deformation produces measurable electrical fields. However, this has received much less scientific attention than electrically induced deformation (Sawahata et al. 1995). This is somewhat surprising, since polyelectrolyte hydrogel based sensors offer unique properties that are unmatched by transducers based on piezoelectric crystalline materials: they are flexible, easy to produce in any desired shape, and, depending on their composition, potentially biocompatible.

In the present contribution, we will summarize our recent work on polyelectrolyte hydrogels as sensor materials. These activities have been along two main lines: on the one hand, we have systematically studied the structure-property relationships, with the goal to identify the physical principles that allow to maximize the voltage produced for a given change in pressure. This required the development of a standardized test method, which exposes the hydrogel sample to a well-defined pressure distribution. On the other hand, we have started to explore applications of hydrogel-based pressure sensors in the context of microfluidic devices.

The remainder of this paper is organized as follows: The next section gives a brief overview of available theoretical approaches to electromechanical effects in hydrogels. We then discuss a test method developed in our group, and compare its results with theoretical predictions and with conventional streaming potential measurements. Finally, we discuss early results on the performance of microfluidic devices exploiting the electromechanical coupling effect for pressure sensing.

2. Theory

Figure 1 shows a conceptual rendition of a polyelectrolyte hydrogel network swollen in water. The ionizable groups are dissociated to an extent that depends on their pKa value in relation to the pH of the solution. Typical acid dissociation constants for carboxylic groups range from from pKa = 4.25 to 6.76 (Park and Robinson 1987; Skouri et al. 1995; Gong et al. 2007).

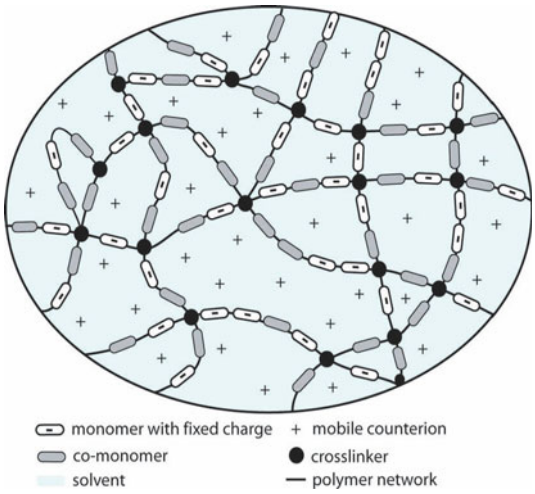


Fig. 1 Concept drawing of a polyelectrolyte hydrogel

If the dry, cross-linked polymer is brought in contact with a solvent, it swells until the gain in free energy due to solvation is balanced by the free energy of stretching of the polymer chains. Since the degree of dissociation of the ions can be very sensitive to temperature, pressure, and pH, the equilibrium swelling volume of polyelectrolyte hydrogels is often a strong function of these external factors (Beltran et al. 1991, De et al. 2002). The swelling behavior of polyelectrolyte gels has been treated in detail on the basis of de Gennes' polymer scaling theory by Rubinstein et al. (1996).

The electromechanical coupling in charged polymers has been approached from two distinct theoretical standpoints: equilibrium and non-equilibrium thermodynamics. The equilibrium picture is based on the Donnan potential: In contact with a polar solvent, the free counterions diffuse out beyond the boundary of the swollen gel until the corresponding gain in entropy is balanced by the resulting electric potential difference between the inside and the outside of the gel (Donnan 1934). This leads to the formation of a charge double layer at the gel/liquid boundary. The magnitude of the Donnan potential depends on the density of dissociable groups in the gel, and the osmotic pressure of the solvent. Applying pressure to the gel is equivalent to a local increase in the osmotic pressure of the solvent, which leads to a measurable local change in the Donnan Potential.

An alternative viewpoint, based on non-equilibrium thermodynamics, has originally been suggested by de Gennes et al. (2000) in the context of the Nafion-type polymer actuators mentioned in the introduction: The electromechanical effect is the result of the coupling between ion flux, driven by the electrical potential difference, and solvent flux, driven by the osmotic pressure gradient. As will be detailed in the following, each viewpoint captures certain aspects of our experimental data on electromechanical coupling in hydrogels.

We have quantified the electromechanical coupling on polyelectrolyte hydrogels using two different experimental approaches. On the one hand, we have measured the streaming potential through a hydrogel membrane. This involves exposing a gel layer, supported between two sieve plates, to a differential solvent pressure, and measuring the resulting electrical potential difference between the upstream and downstream liquid. As discussed in the next section, the variation of the streaming potential with gel composition and salinity of the solvent was very well described by the non-equilibrium thermodynamics approach. On the other hand, we have developed a test method based on indentation of a flat gel sample with a spherical indenter while monitoring the electrical potential distribution. The variation of the coupling with gel composition was found to be proportional to the Donnan potential in the gel, suggesting that the equilibrium picture is appropriate for this type of experiment.

3. Measurement of electromechanical coupling coefficients

3.1 Streaming potential measurements

The streaming potential is defined as the electrical potential difference resulting from the pressure-driven flow of a liquid through a capillary or a membrane. We wanted to test whether the streaming potential and the electromechanical coupling effect in polyelectrolyte hydrogels were related. In order to accomplish this, the variation of the streaming potential coefficient (upstream-downstream voltage difference per unit pressure difference) with the composition of the gel was compared to the variation of the electromechanical coupling constant. We conducted a study of the streaming potential in poly(acrylic acid)/poly(acrylamide) hydrogels. Both the acrylic acid (AA) to acrylamide (AM) ratio and the cross-linking density were varied. In the presence of a buffer at pH = 5, changing the AA content amounts to a variation of the spatial density of negative charges on the polymer backbone, since only the acrylic acid groups dissociate at this pH. In addition, the effect of changing the salinity of the equilibrating solution was measured. Figure 2 shows the measurement apparatus. Two chambers containing a buffer solution of pH 5.00 and defined salinity were separated by the polyelectrolyte hydrogel sample, which was held in place by a couple of sieve plates. Ag/AgCl electrodes were used to measure the electrical potential difference between the two solvent chambers. A time-modulated pressure difference was imposed using an electronic pressure controller. Compressed nitrogen gas was used as the primary pressure source.

The concentration of ionizable groups affects the swelling behavior of the gels. Since more highly charged gels tend to swell more,

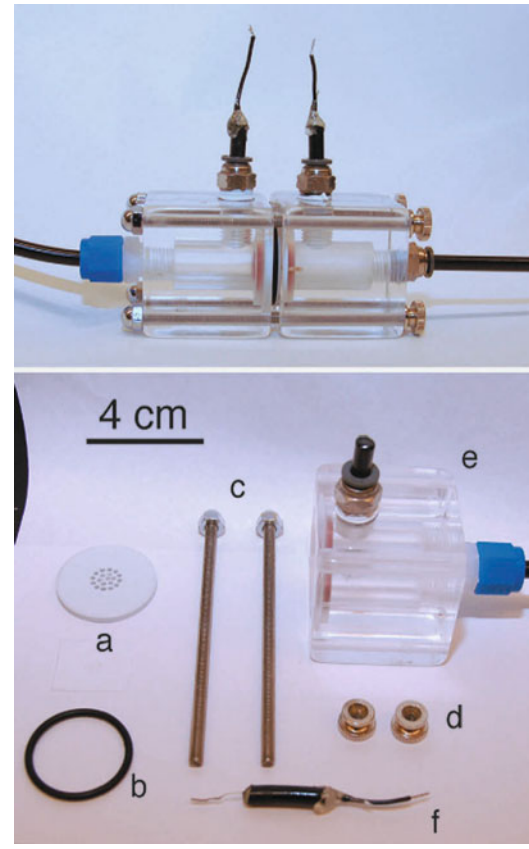


Fig. 2 Pressure cell assembly. **a** sieve plate, **b** o-ring, **c** steel rods to connect half-cells, **d** cell-tightening screws, **e** acrylic block cell, **f** electrode

the final charge density in the equilibrium swelling state is a non-linear function of gel composition, as shown in Fig. 3 B.

Accurate measurements of differential voltages in electrochemical settings require measurement amplifiers with high input impedance, typically in excess of $10^{14}\,\Omega$. Due to unavoidable stray capacitances in the measurement setup, this leads to spurious, slowly drifting signal readouts. In order to eliminate these, we used a slow sinusoidal modulation of the input pressure. This allowed for the removal of the ultra-low frequency components in the data by digital high-pass filtering. Figure 4 shows a typical pressure input profile along with the streaming potential

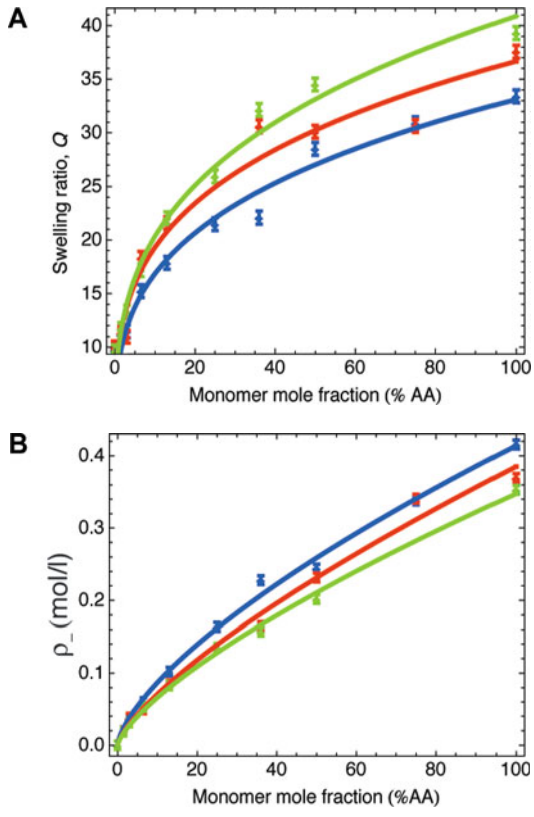


Fig. 3 **A** Swelling ratio of acrylic acid/acrylamide gels as a function of gel composition. The three different traces represent different degrees of cross-linking (cross-linker concentrations 0.0079M, 0.014M, and 0.032M), with the least cross-linked gel swelling the most. **B** Spatial densities of ionizable groups in the same gels after swelling

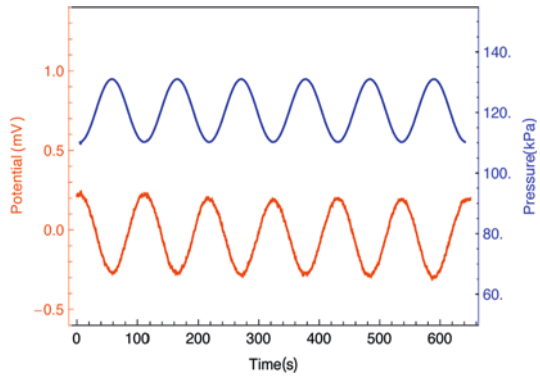


Fig. 4 Pressure difference input (blue, top), and resulting voltage readout signal (red, bottom) from a typical streaming potential measurement. Slow drift in the voltage response has been removed by digital high-pass filtering

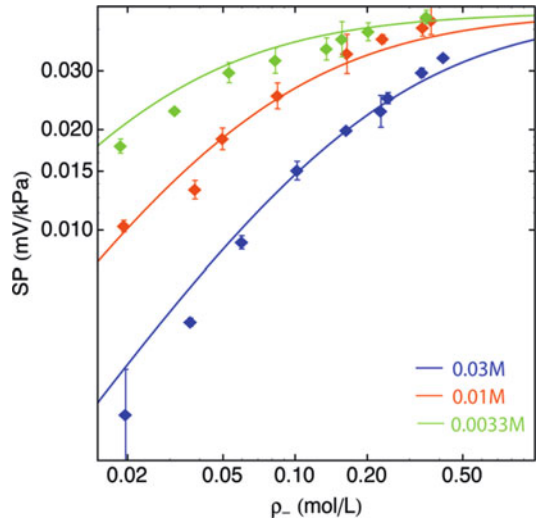


Fig. 5 Streaming potential coefficient for different gels as a function of solvent salinity (indicated as concentration of NaCl in the buffer). Low salinity leads to a large response. The lines represent a simple theory based on de Gennes' non-equilibrium thermodynamics treatment

response (Fiumefreddo and Utz, 2010). The voltage output is a linear function of the input pressure, as shown by Fourier analysis: No additional harmonics were detectable at a signal/noise ratio in excess of 30 dB. The proportionality coefficient relating the pressure amplitude to the observed streaming potential is shown as a function of solution salt concentration, and density of dissociable groups, in Fig. 5. The streaming potential increases with charge density, and depends on the salinity of the solvent. This is shown in Fig. 5: Higher salt concentrations generally lead to smaller streaming potentials. This makes immediate sense due to the screening effects of the additional charged ions. Figure 5 also shows that the solvent salinity affects the dependence of the streaming potential on the intrinsic charge density of the gel. A simple theory based on de Gennes'

non-equilibrium thermodynamic treatment (Fiumefreddo and Utz 2010) can explain this behavior quantitatively, as shown by the solid lines in Fig. 5. The predictions are calculated with a single adjustable parameter: the degree of dissociation of the ionizable groups. In the present context, a value of about 30 % dissociation gives a good fit to the data. This is a reasonable value, corroborated by the known properties of polyacrylic acid (Park and Robinson 1987; Skouri et al. 1995, Gong et al. 2007), as well as independent observations from our own work (Prudnikova and Utz 2010). The streaming potential increases with the charge density in the gel, but reaches a saturation point when the density of negative charges on the polymer backbone becomes comparable to the concentration of free negative salt ions.

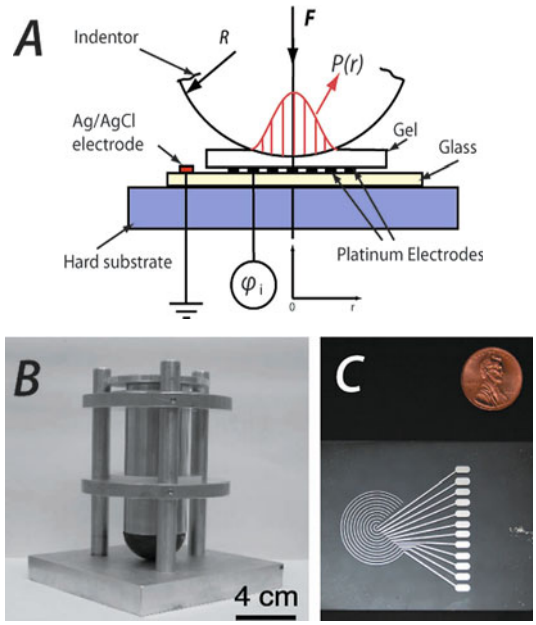


Fig. 6 **A** Principle of the indentation experiment. **B** Indenter assembly. **C** Pt electrode pattern made using a mask lift-off microfabrication technique

3.2 Indentation testing

With the goal of developing a standardized, systematic method to test the response of polyelectrolyte gels to inhomogeneous pressure distributions, we have designed a test based on indentation of a flat polymer gel sample with a spherical indenter. The principle of the method is shown in Fig. 6. The sample is placed on top of a concentric ring pattern of microfabricated Pt electrodes with diameters from 2 mm to 24 mm. The sample is indented from above by a sphere of 20 mm radius. The indenter force is produced by a pneumatic cylinder (not shown), which is in turn connected to a pressurized gas source via an electronic pressure controller (MKS Instruments, Inc.). The load was slowly modulated in time in order to eliminate problems with slow signal drift. The voltages picked up by the platinum electrodes were amplified using a home-built high-impedance amplifier circuit, which was connected to a National Instruments A/D converter unit. For full details on the experimental setup and procedure see Prudnikova and Utz (2010).

The elasticity problem of a spherical body in contact with a flat surface has been treated by Hertz (1882). In the present case, the indented sample does not represent a homogeneous half-space, but a relatively thin, soft sample layer (approx 1–2 mm) on top of a very stiff support. This arrangement is described by a modified version of the original Hertzian theory (Johnson 1987). Figure 7 shows a typical data set obtained from this experiment. The signal is recorded simultaneously from all electrodes, which provides spatial resolution. As shown in Fig. 7 C, the resulting amplitude profile follows the shape of the modified Hertzian contact pressure. The electromechanical coupling constant can therefore be determined from the known input force amplitude by a straightforward fit to the spatial amplitude profile.

Gels with systematically varying com-

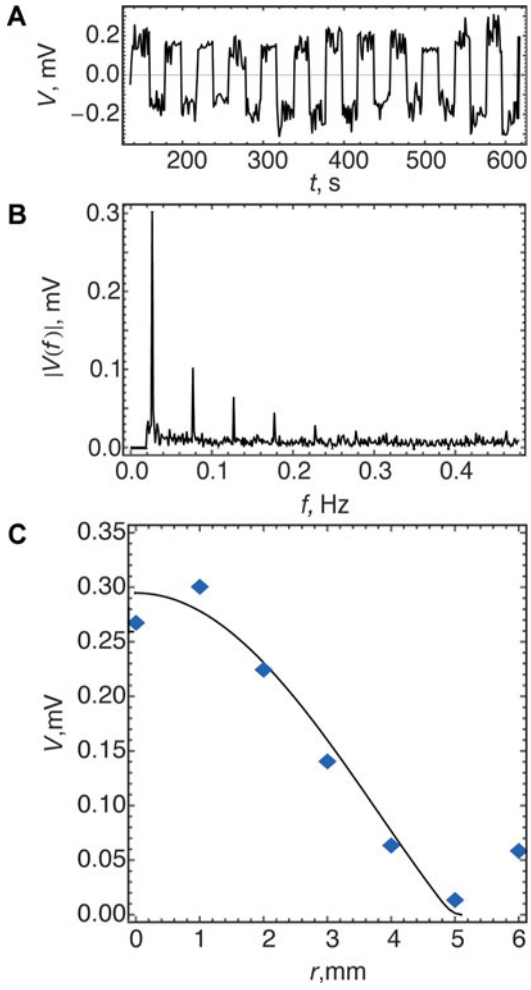


Fig. 7 **A** Voltage output from a single electrode. A rectangular force signal with a period of 40s has been applied to the indenter. **B** Absolute value spectrum of the signal shown in A. The odd harmonics correspond closely to a square wave, confirming the linearity of the voltage as a function of pressure. **C** Spatial profile of the voltage amplitude as a function of radial distance from the center of the indenter. The distribution follows the pattern expected from the modified Hertzian contact model (solid line)

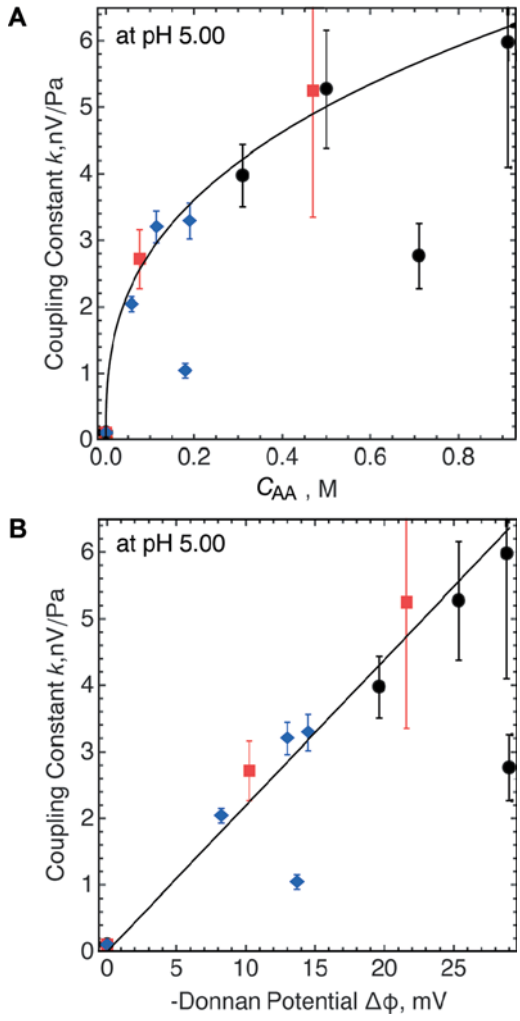


Fig. 8 **A** Electromechanical coupling constant measured by indentation for different gels as a function of spatial density of ionizable groups. The square, round, and diamond-shaped symbols represent different cross-link densities. **B** The same data plotted as a function of the independently measured Donnan potential of each gel

position were used for these experiments. Their cross-link densities, however, were somewhat different from those used for the streaming potential measurements. This was necessary because the samples need to be slightly more robust for the indentation experiment. Conversely, the samples used for the indentation experiments were too brittle for the streaming potential experiments, and would break upon clamping inside the sample holder. While the results from the streaming potential and indentation measurements cannot be compared directly as a result, it is still significant that the streaming

potential and the electromechanical coupling coefficients are similar in magnitude, and behave very alike upon variation of the acrylic acid content in the sample.

Indentation results are compiled in Fig. 8. Similar to the streaming potential, the electromechanical coupling coefficient increases with the density of ionizable groups in the sample. However, in contrast to the streaming potential coefficients, the electromechanical coupling constant seems to be independent of the cross-link density; it is purely a function of the density of ionizable groups in the gel, as evidenced by the fact that the samples with different cross-link densities fall on the same curve in Fig. 8A. Figure 8B plots the same data, but instead of the density of ionizable groups, the Donnan potential, i. e., the equilibrium electrical potential difference between the inside of the gel and the surrounding solvent, is used as horizontal axis (Prudnikova and Utz 2010). As is obvious from the figure, the electromechanical coupling coefficient is proportional to the Donnan potential.

The data presented here suggests that while the exposure of the gel to a gradient in pressure of the surrounding solvent (streaming potential) is better described by non-equilibrium thermodynamics and transport theory, the response to mechanical pressure directly applied to the gel seems to be adequately accounted for by the notion of a simple modulation of the Donnan potential.

4. Pressure sensors based on hydrogels

As an example of a hydrogel-based pressure and touch sensor, the device shown in Fig. 9 was built. It consists of a pattern of 17 circular Au electrodes. The electrodes are separated radially by 1.75 mm. The access leads are protected by a 4 µm layer of photoresist, exposing only the circular tip of each electrode to the polyelectrolyte gel. The gel is produced by casting the monomer solution between glass slides held apart by 250 µm spacers, subsequent polymerization by UV irradiation, and equilibration at pH 5, resulting in samples of 300 µm thickness. Finally, a 50 µm layer of poly(dimehtyl siloxane) (PDMS) protects the gel from drying out.

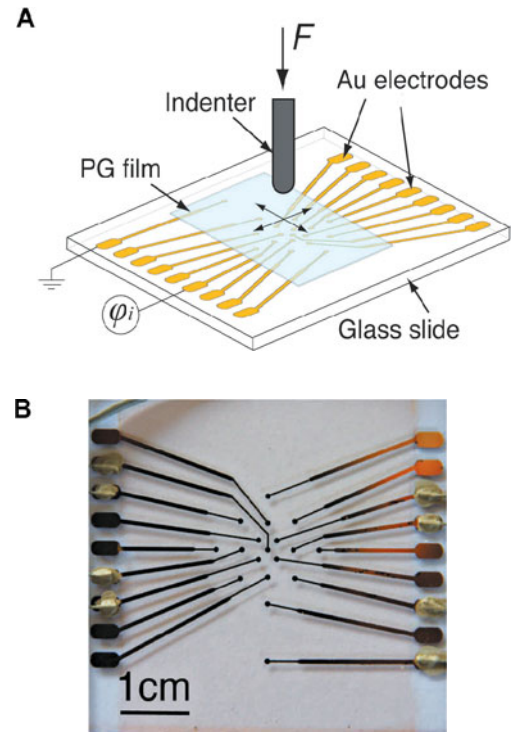


Fig. 9 **A** Microfabricated pressure/touch sensor based on polyelectrolyte hydrogel layer (PG film). 17 electrodes are patterned onto a glass support. **B** Micrograph of the Au electrode pattern

This sensor was exposed to loading with a spherical indenter of 2.15 cm radius, at a force oscillating between 2 and 3 N. In combination with the modulus of the gel layer (0.55 MPa), this produces a contact radius in the vicinity of 2 mm, comparable to the separation of the electrodes.

The results are shown in Fig. 10A; the sig-

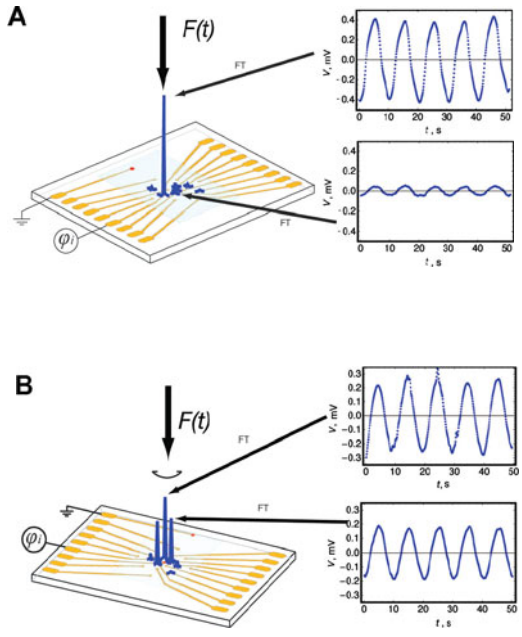


Fig. 10 **A** Signals obtained by loading the microfabricated touch sensor with a sharp tip radius: only a single electrode produces a large signal. **B** Loading with a blunt object leads to excitation of multiple nearby electrodes

nal from the electrode closest to the center of the contact area produces a large signal, whereas nearby electrodes exhibit amplitudes more than a factor of 10 smaller. By contrast, indentation with a sphere of 6.5 cm radius produces the picture shown in Fig. 10 B: now, the signals from neighboring electrodes are of comparable strength.

This simple demonstration device shows that it is possible to build a sensor that can provide quantitative force information, spatial resolution, as well as the distinction between sharp and blunt contact, by simply placing a polyelectrolyte hydrogel layer on top of a suitably laid out pattern of electrodes.

A potential application of hydrogel-based pressure sensors lies in the field of microfluidic devices. The main idea is to separate the fluidic network from a layer of polyelectrolyte gel by a thin, flexible membrane of PDMS. In this way, the pressure distribution inside the network is transmitted into the gel layer, which then converts it into an electric potential distribution. The latter can be measured at any desired point if electrodes are patterned into the glass support of the gel layer. This is shown schematically in Fig. 11 A: a linear fluidic channel is instrumented with two electrodes, at either end. The voltage difference between the electrodes is proportional to the pressure drop

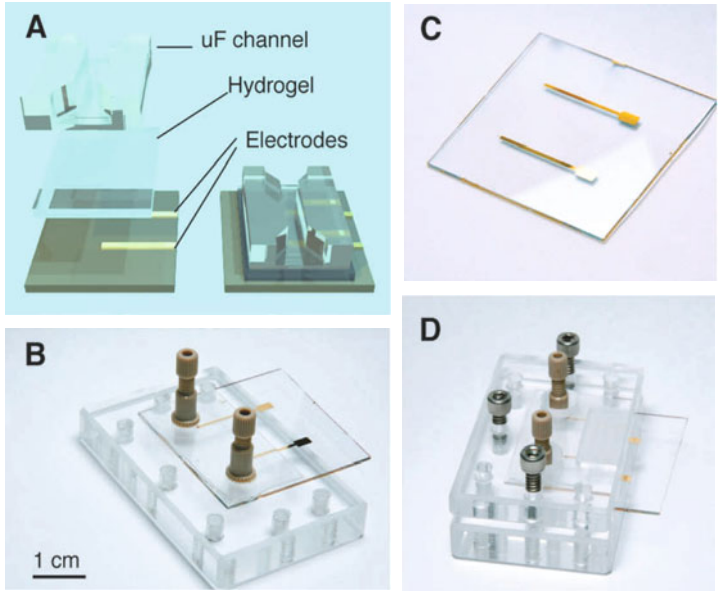


Fig. 11 **A** Schematic of a hydrogel-based flow sensor in a µF-channel - electrodes are patterned onto the support layer and are insulated exposing only the area beneath the channel to the contact with a polymer gel. The pressure drop across the channel translates into a potential difference measured across two electrodes. **B** µF-channel with electrodes and fluidic connectors on plastic holder. **C** Microfabricated Au electrodes **D** complete experimental assembly

across the channel. If the viscosity of the fluid is known, this can be used to quantify the rate of flow.

We have built such a device, shown in Fig. 11B–D. A channel of 24 mm length, 86 µm deep, and 190µm wide was prepared via conventional wet etching in a glass substrate, and covered with 200 µm PDMS layer. The electrodes (Fig. 11C) were patterned on a glass support by sputtering and photoresist lift-off. Only their circular tips, measuring 200 µm accross, are exposed to the polyelectrolyte gel (PG) film, the rest of the metal structure was protected by a layer of photoresist. The electrodes are separated along the channel by 20 mm. The sandwich structure of the fluidic channel, hydrogel layer, PDMS layer, and electrodes was assembled and aligned by hand under a stereo microscope, and held together on a plastic support (Fig. 11D).

Water was forced through the fluidic channel by applying a sinusoidal time-pressure profile to a water reservoir using an electronic pressure controller. This lead to a sinusoidally varying flow rate between 2 and 5 µL/s, with a resulting voltage amplitude of about 2 mV. These data are as yet preliminary, and reflect some unresolved imperfections of the chip. In particular, there asymmetry of the signal (positive vs negative peaks), which seems linked to gasket problems. However, the data confirms the original concept of a pressure-instrumented microfluidic network.

Conclusions and outlook

In summary, a quantitative experimental study of the electromechanical properties of polyelectrolyte hydrogels based on poly(acrylic acid) and poly(acryl amide) has been presented. The electromechanical coefficient was found to follow the Donnan potential of the gels. The streaming potential, which describes the electrical response to the forced flow of solvent through the gel, was also measured. While the streaming potential and the electromechanical voltage are of similar magnitude, their scaling with the crosslink density of the gel was found to differ. While the electromechanical effect is invariant of the crosslink density (at constant spatial charge density in the gel), the streaming potential is higher for more highly crosslinked systems.

We have demonstrated the use of polyelectrolyte hydrogel layers in microfluidic systems in order to implement pressure sensors. Two simple devices have been built: a spatially resolved touch sensor, which can distinguish blunt from sharp objects, and a microfluidic flow meter based on the viscous pressure drop across a channel of known fluidic resistance.

Future efforts in this field will be devoted to consolidate the as-yet preliminary data on the touch sensor and the flow meter. In addition, the experimental results on the relationship between the gel structure, state of stress, and resulting electrical potential needs to be integrated with recent continuum models of gel mechanics (Hong et al. 2008). Such a coupled electromechanical diffusion- and deformation theory will allow large-scale finite element modeling (FEM) of complex polyelectrolyte gel sensor and actuator systems, enabling the design of functional devices.

Acknowledgments

This work has been supported by the US National Science Foundation. The streaming potential experiments were performed by Ms. Andrea Fiumefreddo as part of her Master of Science thesis. We also gratefully acknowledge helpful discussions with Prof. Andrey Dobrynin.

References

Bassetti M, Chatterjee A, Aluru N, Beebe D (2005) Development and modeling of electrically triggered hydrogels for microfluidic applications. J Microelectromech Syst 14: 1198–1207

Beebe D, Moore J, Bauer J, Yu Q, Liu R, Devadoss C, Jo B (2000) Functional hydrogels structures for autonomous flow control inside microfluidic channels. Nature 404: 588–590

Beltran S, Baker J, Hooper H, Blanch H, Prausnitz J (1991) Swelling equilibria for weakly ionizable

temperature sensitive hydrogels. Macromolecules 24: 549–551

De S, Aluru N, Johnson B, Crone W, Beebe D, Moore J (2002) Equilibrium swelling and kinetics of pH-responsive hydrogels: models, experiments, and simulations. J Microelectromech Syst 11: 544–555

Doi M, Matsumoto M, Hirose Y (1992) Deformation of ionic polymer gels by electric fields. Macromolecules 25: 5504–5511

Donnan FG (1934). Z Phys Chem (Munich) 168: 369–380

Fiumefreddo A, Utz M (2010) Streaming potential in poly(acrylic acid)/poly(acrylamide) hydrogels. Manuscript in preparation

de Gennes PG, Okumura K, Shahinpoor M, Kim KJ (2002) Mechanoelectric effects in ionic gels. Europhys Lett 50: 513–518

Gong P, Wu T, Genzer J, Szleifer I (2007) Behavior of surface-anchored poly(acrylic acid) brushes with grafting density gradients on solid substrates: 2. Theory. Macromolecules 40: 8765–8773

Hertz H (1882) Verh Ver Beförderung Gewerbefleisses 61: 449–462

Hong W, Zhao X, Zhou J, and Suo, Z (2008), A theory of coupled diffusion and large deformation in polymeric gels, J Mech Phys Solids 56: 1779–1793

Johnson KJ (1987) Contact mechanics. Cambridge University Press, Cambridge, UK,pp 138–141

Nemat-Nasser S (2002) Micromechanics of actuation of ionic polymer-metal composites. J Appl Phys 92: 2889–2915

Park H, Robinson JR (1987) Mechanisms of mucoadhesion of poly(acrylic acid) hydrogels. Pharm Res 4: 457–464

Prudnikova K, Utz M (2010) Electromechanical characterization of polyelectrolyte gels by indentation. Macromolecules 43: 511–517

Richter A, Kuckling D, Howitz S, Gehring T, Arndt K (2003) Electronically controllable microvalves based on smart hydrogels: magnitudes and potential applications. J Microelectromech Syst 12: 748–753

Rubinstein M, Colby R, Dobrynin A, Joanny J (1996) Elastic modulus and equilibrium swelling of polyelectrolyte gels. Macromolecules 29: 398–406

Sawahata K, Gong JP, Osada Y (1995) Soft and wet touch-sensing system made of hydrogel. Macromol Rapid Commun 16: 713–716

Shahinpoor M, Bar-Cohen Y, Simpson JO, Smith J (1998) Ionic polymer-metal composites (IPMCs) as biomimetic sensors, actuators and artificial muscles – a review. Smart Mater Struct 7: R15-R30

Shahinpoor M, Kim KJ (2002) Solid-state soft actuator exhibiting large electromechanical effect. Appl Phys Lett 80: 3445–3447

Skouri R, Schosseler F, Munch J, Candau S (1995) Swelling and elastic properties of polyelectrolyte gels. Macromolecules 28: 197–210

Tanaka T, Nishio I, Sun S, Ueno-Nishio S (1982) Collapse of gels in an electric field. Science 218: 467–469

Wallmersperger T, Kroplin B, Gulch RW (2004) Coupled chemo-electro-mechanical formulation for ionic polymer gels – numerical and experimental investigations. Mech Mater 36: 411–420

Yamaue T, Mukai H, Asaka K, Doi M (2005) Electrostress diffusion coupling model for polyelectrolyte gels. Macromolecules 38: 1349–1356

Yao L, Krause S (2003) Electromechanical responses of strong acid polymer gels in DC electric fields. Macromolecules 36: 2055–2065

Single-molecule detection of proteins using nanopores

25

Liviu Movileanu

Contents

Abstract

Recently, protein and synthetic nanopores have been employed extensively as single-molecule probes to illuminate the functional features of proteins, including their binding affinity to different ligands, backbone flexibility, enzymatic activity and folding state. In this chapter, I present a brief overview in this emerging area of biosensing. The underlying principle of detection is that the device is based upon a single nanopore drilled into an insulating membrane, which is immersed in a symmetric chamber containing electrolyte solution. The application of a transmembrane potential across the membrane will enable the recording of a well-defined electric current due to the flow of ions crossing the nanopore. The partitioning of single proteins into the interior of the nanopore is detected by discrete current fluctuations that depend upon the interaction between the proteins and the nanopore. The detection mechanisms include chemical modification and genetic engineering of protein nanopores, electrophoretic capture of proteins via movable nucleic acid arms, and functionalization of the inner surface of synthetic nanopores. This approach holds promise for the exploration of proteins at high temporal and spatial resolution. Moreover, nanopore probe techniques can be employed in high-throughput devices used in biomedical molecular diagnosis and environmental monitoring.

Liviu Movileanu
Syracuse University, Department of Physics
201 Physics Building
Syracuse, NY 13244-1130, USA
e-mail: lmovilea@physics.syr.edu

1. Introduction

There is an unprecedented need for novel methodologies to detect a broad range of substances, from small organic molecules to macromolecules, bacteria and viruses. In this chapter, we discuss the power, versatility

and tractability of the nanopore technique as a single-molecule probe that holds potential to reveal several important features of polypeptides and proteins. The underlying principle of nanopore probe techniques is simple: the application of a voltage bias across an electrically insulated membrane enables the measurement of a tiny picoamp-scale transmembrane current through a single nanopore. Each molecule, translocating through the nanopore, produces a distinctive current blockade, the nature of which depends on its biophysical properties as well as the molecule-nanopore interaction. Such an approach proves to be quite powerful, because single small molecules and biopolymers are examined at very high spatial and temporal resolutions (Bayley 1999; Bayley et al. 2000, 2004; Bayley and Martin 2000; Bezrukov 2000; Deamer and Akeson 2000; Bayley and Cremer 2001; Deamer and Branton 2002; Meller 2003; Bayley and Jayasinghe 2004; Dekker 2007; Martin and Siwy 2007; Howorka and Siwy 2008; Howorka and Siwy 2009; Movileanu 2008, 2009). Here, we also discuss recent studies that provided a mechanistic understanding of the forces that drive protein translocation through a nanopore. These measurements facilitate the detection and exploration of the conformational fluctuations of single molecules and the energetic requirements for their transition from one state to another. We present recent strategies for engineering new functional nanopores, in organic and silicon-based materials, and with properties that are not encountered in nature. From a practical point of view, this methodology shows promise for the integration of engineered nanopores into nanofluidic devices, which would provide a new generation of research tools in nanomedicine and high-throughput instruments for molecular biomedical diagnosis and environmental monitoring.

Stochastic biosensors represent a unique class of single-molecule detectors that is based upon the alteration of electrical current by distinct molecules that interact either transiently or permanently with an engineered recognition group located within a nanopore (Martin and Siwy 2007; Howorka and Siwy 2008; Movileanu 2009) (Fig. 1A). Each transient interaction of a single molecule with the engineered recognition group produces a well-defined current blockade, the duration of which depends upon several parameters, including the size, charge and shape of the interacting molecule, the accessibility of the engineered recognition group (e. g., the diameter of the inner surface of the nanopore; the position of the recognition group), and the strength of the analyte-reactive group interaction (Fig. 1B) (Bayley and Martin 2000; Bayley and Cremer 2001). This current blockade occurs as a result of the interacting molecule-induced obstruction of the ionic flow through the nanopore. Therefore, the partitioning of a single molecule into the interior of the nanopore is probed by the resistive-pulse technique in real time (Bezrukov 2000).

In the past decade, research in the area of nanopores has been stimulated by the keen interest of this technique in single-molecule sequencing of nucleic acids (Kasianowicz et al. 1996; Akeson et al. 1999; Henrickson et al. 2000; Meller et al. 2000; Howorka et al. 2001a,b; Howorka and Bayley 2002; Kasianowicz 2004; Meller et al. 2005; Bayley 2006; Kim et al. 2006; Butler et al. 2007; Rhee and Burns 2007; Branton et al. 2008; Butler et al. 2008; Howorka and Siwy 2008; McNally et al. 2008; Wanunu et al. 2008 Comer et al. 2009; Dorvel et al. 2009). Despite numerous obvious challenges in using nanopores for nucleic acid sequencing (Branton et al. 2008), recent studies have demonstrated that nanopores represent versatile single-molecule probes for a broad range of molecules and their assemblies (Dekker 2007; Griffiths 2008; Hayden 2008). An extensive research on biosensing has been carried out with the

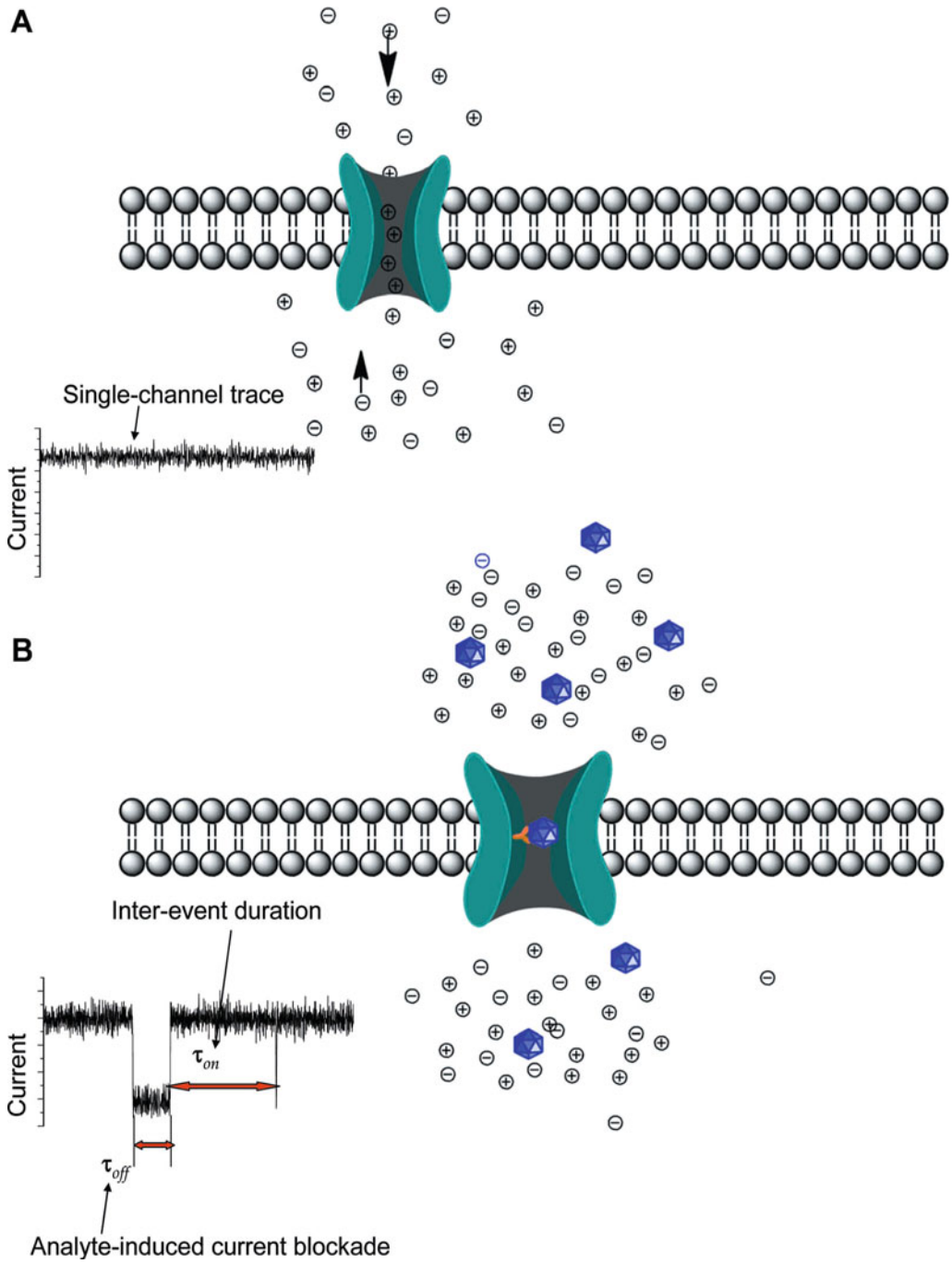


Fig. 1 Principle of stochastic sensing with a permanently open protein nanopore. An open protein nanopore represents the permeation pathway for transmembrane transport of small ions. **A** When a transmembrane potential is applied, a single-channel electrical current is readily recorded. **B** The nanopore interior is either genetically engineered or chemically modified by covalent attachment of a functional reactive group. Analytes are added to one of the chambers and transiently interact with the functional reactive group located within the pore interior. This transient interaction results in a well-defined single-channel current blockade, the nature of which depends on the strength of the interaction between the analyte molecule and the functional reactive group

staphylococcal α-hemolysin (αHL) protein. αHL forms a mushroom-shaped heptameric transmembrane pore in planar lipid bilayers (Fig. 2 A) (Song et al. 1996). This protein is a highly tractable β-barrel-type pore because of the following unique combination of attributes: (i) the availability of its high-resolution crystal structure (Song et al. 1996), (ii) the ease of its genetic engineering (Bayley and Cremer 2001; Movileanu et al. 2001), (iii) its large single-channel conductance (Bezrukov et al. 1996; Kasianowicz et al. 1996), which facilitates high time-resolution single-channel electrical recordings, (iv) its high thermal stability (Kang et al. 2005), (v) the pore remains open indefinitely under a wide range of experimental conditions, such as pH (Bezrukov and Kasianowicz 1993; Kasianowicz and Bezrukov 1995), ionic strength (Krasilnikov et al. 2006; Rodrigues et al. 2008), temperature (Kang et al. 2005; Jung et al. 2006), transmembrane potential (Movileanu et al. 2005) and osmolytes (Bezrukov et al. 1996; Krasilnikov and Bezrukov 2004). These attributes made the αHL protein a suitable nanopore platform for single-molecule stochastic sensing of small molecules and large biomolecules (Bayley and Martin 2000; Bayley and Cremer 2001; Howorka et al. 2001 b; Bayley 2006; Rhee and Burns 2007; Branton et al. 2008; Griffiths 2008; Hayden 2008; Howorka and Siwy 2008, 2009; Movileanu 2008).

2. Watching polypeptide translocation

In the last decade, several groups have extensively examined the translocation of polypeptides through protein pores using single-channel electrical recordings. These studies, which employed cyclic petides (Sanchez-Quesada et al. 2000), and α-helical (Movileanu et al. 2005; Sutherland et al. 2005; Stefureac et al. 2006) and β-hairpin (Goodrich et al. 2007) polypeptides, revealed specific kinetic signatures of polypeptide-pore interactions. These kinetic signatures depended upon biophysical features of the translocating polypeptides, such as their charge, length, folding and stability. In general, the single-channel electrical recordings showed very short-lived current blockades assigned to bumpings of the polypeptide into the pore entrance and long-lived current blockades assigned to major partitionings of the polypeptide into the interior of the pore. One simple model of the interaction between a polypeptide and a protein pore can be illustrated by a two-barrier, single-well free-energy lansdcape (Movileanu et al. 2005; Wolfe et al. 2007; Movileanu 2008). This kinetic scheme was observed with α-helical peptides (Movileanu et al. 2005; Wolfe et al. 2007). In contrast, other studies showed different kinetic schemes. For example, in a recent study, we combined single-molecule electrical recordings and Langevin dynamics simulations to explore the partitioning of β-hairpin peptides into the αHL protein pore (Fig. 2) (Goodrich et al. 2007). The dwell time derived from event histograms showed a single exponential decay-dependence on the applied transmembrane potential (Fig. 3). This finding suggested that the β-hairpin-pore interaction undergoes a single-barrier free energy lanscape. Moreover, we showed that highly unfolded β-hairpin polypeptides entered the pore in an extended conformation, producing fast single-file translocation events. This result contrasted with that of the translocation of structured β-hairpin polypeptides, which occurred more slowly, producing long-lived current blockades (Goodrich et al. 2007). In Fig. 3, we illustrate that more strutured β-hairpins, K41 and G41, translocate more slowly than the unstructured polypep-

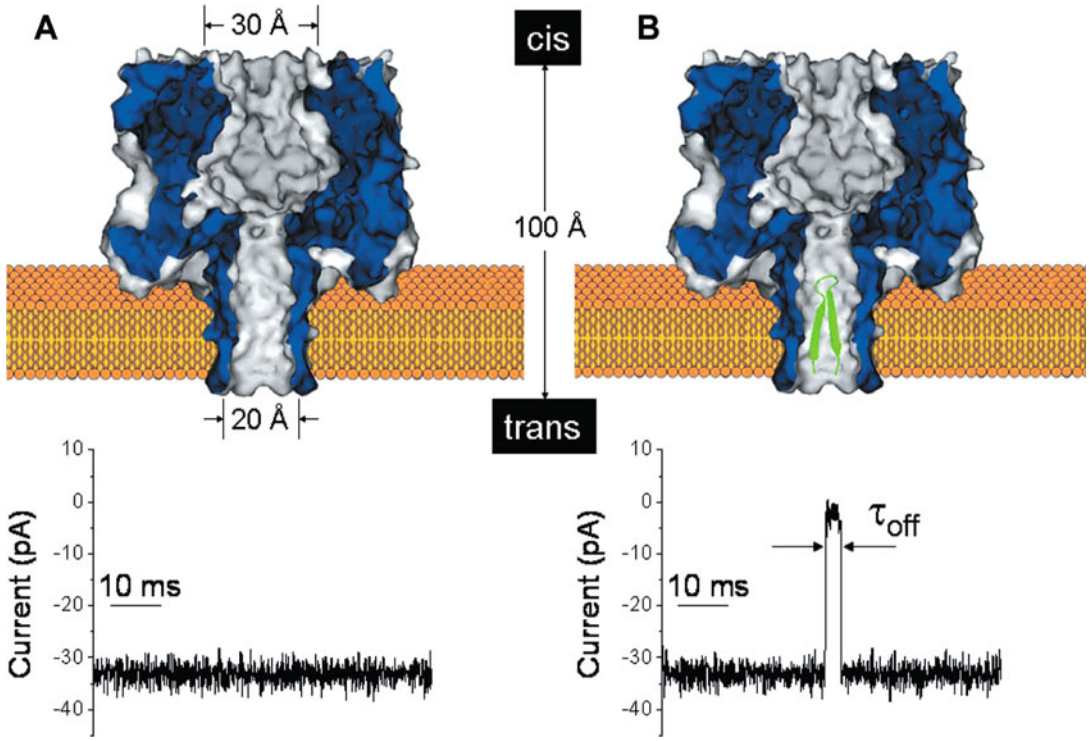


Fig. 2 The interaction of β-hairpin peptides with an aHL pore produces transient current blockades: **A** αHL forms a pore that remains open for long periods. **B** The translocation of a β-hairpin peptide through an αHL pore produces a transient current blockade. Reproduced, with permission, from Goodrich et al. (2007)

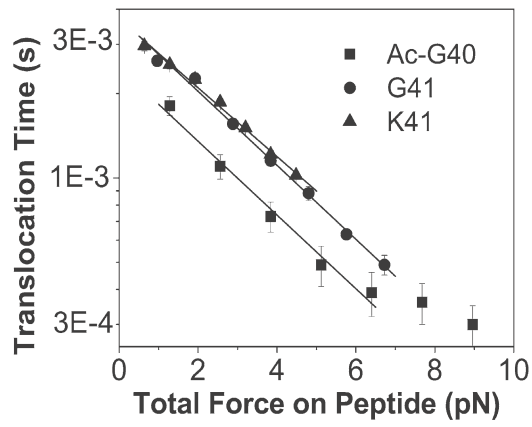


Fig. 3 The dependence of the dwell time of the current blockades produced by the **β**-hairpin peptides on the electric force. The electric force was derived as $F = qV/l$, where the effective charge q is half of the peptide charge, V is the transmembrane potential, and $l = 50$ Å is the length of the β-barrel lumen. Reproduced, with permission, from Goodrich et al. (2007)

tide Ac-G40. This finding is presented by an increased translocation time measured for more structured peptides regardless of the applied total force on peptide (Fig. 3).

Extensive research on polypeptide sensing with nanopores has been pursued by the group of Jeremy Lee (Sutherland et al. 2005; Stefureac et al. 2006, 2008; Stefureac and Lee 2008). They showed that the αHL protein pore can be employed to discriminate several polypeptides based upon their size and conformation (Sutherland et al. 2005; Stefureac et al. 2006). In addition, these studies revealed that the electrical recordings hold potential for discovering the presence of intermediate conformations of the polypeptides, which were not apparent from the circular dichroism (CD) spectra (Stefureac et al. 2008; Stefureac and Lee 2008).

One way to alter the kinetics of the

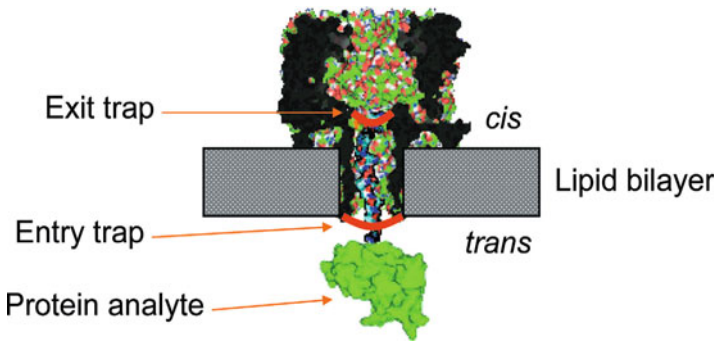


Fig. 4 The αHL protein pore is equipped with electrostatic traps engineered at the entry and exit of the channel. A single electrostatic trap consists of an iris of seven negatively charged aspartic acid residues. Reproduced, with permission, from Mohammad et al. (2008)

polypeptide-pore interactions is to engineer attractive or repulsive groups within the pore interior. These engineered groups might impact the free energy lansdcape of the polypeptide-pore interactions. Recently, we were able to engineer attractive groups, simply called traps, within the interior of the αHL protein pore. Two irises of negatively charged traps have been engineered at the entry and exit of the β-barrel part of the αHL protein pore (Fig. 4) (Wolfe et al. 2007). The engineered traps consisted of aspartic acid residues. In a first study, we examined the translocation of short positively charged polypeptides (~25 residues in length) through αHL protein pores equipped with engineered traps (Wolfe et al. 2007). The label-free single-channel recordings enabled the simultaneous determination of the rate constants of association (k_{on}) and dissociation (k_{off}) (Mohammad et al. 2008; Mohammad and Movileanu 2008). k_{on} is defined as $1/(c\tau_{on})$, where τ_{on} is the inter-event interval, and c is the concentration of the substrate in the bilayer chamber. k_{off} is defined as $1/\tau_{off}$, where τ_{off} is the mean blocked time of the transient substrate-induced current blockades (Fig. 1B). The dissociation constant is given by $K_d = k_{off}/k_{on}$, and the association constant is given by $K_a = 1/K_d$, allowing for the calculation of the standard free energy $\Delta G^o = -RT \ln K_a$. The electrostatic trap engineered at the entry of the pore had a major impact on the rate constant of association (Wolfe et al. 2007). By contrast, the rate constant of dissociation was only marginally altered. Surprisingly, when the entry and exit traps were concurrently engineered within the interior of the nanopore, significant increase in the rate constants of association and dissociation were observed. By analogy with ion-conducting protein channels, these negatively charged traps acted as binding sites for the positively charged polypeptide analyte. The results of this study indicated that the two traps together catalyze the net flow of polypeptides from one side of the membrane to the other, as compared with either the wild type-αHL pore or a single trap-containing αHL protein pore.

In a subsequent study, a similar experimental design was used for examining the polypeptide–pore interactions, but with a much larger interacting protein. In this case, the analyte was a fusion protein (pb_2-Ba) that consisted of a positively charged presequence polypeptide of cytochrome b_2 (pb_2) fused to the folded ribonuclease barnase (Fig. 4) (Mohammad et al. 2008). The results were significantly different from those obtained with short polypeptides. For example, the addition of nanomolar concentrations of pb_2-Ba protein to the chamber produced infrequent and short-lived current blockades with the WT-αHL protein pore (Fig. 5A). This finding was interpreted in terms of a large energetic penalty of the folded protein to

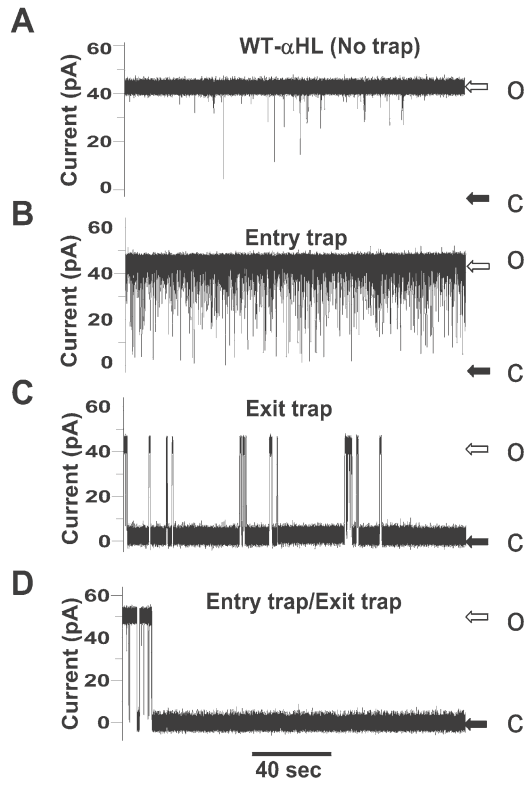


Fig. 5 The effect of the barnase protein on the single-channel electrical recordings performed with the wild-type (WT) and electrostatic trap-containing αHL pores. Representative single-channel electrical recordings with the WT-αHL (**A**), K131 D_7 (**B**), K147 D_7 (**C**) and K131 D_7/K147 D_7 (**D**) pores in the presence of 200 nM pb_2(95)-Ba added to the chamber. The white and black arrows indicate the open and closed states, respectively. Reproduced, with permission, from Mohammad et al. (2008)

partition into the interior of the trap-free αHL protein pore. In contrast, single-channel recordings with an αHL protein pore, which contained a single trap engineered at the entry of the pore, revealed a dramatically increased frequency of the current blockades (Fig. 5B). The event frequency increased by almost two orders of magnitude, as compared with the WT-αHL pore. This finding was consistent with an effective recruitment of the small folded proteins by the engineered electrostatic trap, which was exposed to the aqueous phase. Interestingly, when the electrostatic trap was engineered at the exit of the β-barrel (Fig. 4), away from the aqueous phase, the transient current blockades became long lived, with duration in the range of tens of seconds to minutes (Fig. 5C). The different results from experiments in which the trap was engineered either at the entry or at the exit of the pore illuminated a significantly altered kinetics of the folded protein-pore interactions. Remarkably, the observed transient current blockades became permanent when the pb_2-Ba protein interacted with double trap-containing αHL protein pore (Fig. 5D). These results demonstrate that protein design is instrumental in obtaining major alterations in the single-channel electrical signature as a result of minute changes within the interior of the pore.

Using a similar approach, Guan and colleagues have designed αHL protein pores with hydrophobic traps engineered within the interior of the pore (Zhao et al. 2009b). They studied the translocation of polypeptides containing hydrophobic aminoacids through such hydrophobic traps-containing αHL protein pores. As expected, they found increased interactions between hydrophobic residues-containing peptides and a trap-containing αHL protein pore. Increasing lengths of the translocating peptide produced decreased rate constants of association and dissociation. On the other hand, additional engineering of the αHL protein pore, in the form of many aromatic binding sites, resulted in a stronger binding affinity between the translocating peptide and the hydrophobic trap-containing αHL protein pore. This study is also important, because Guan and colleagues were able to discriminate among polypeptides that differed by a single residue. Therefore, it was suggested that the stochastic biosensing with engineered protein pores holds great promise for peptide and protein sequencing.

3. Probing protein-ligand interactions

Probably, the most important limitation of polypeptide translocation through nanopores is the inability to capture all kinetic information of the polypeptides in the aqueous phase. This limitation has been overcome by attaching a flexible linker, a poly(ethylene glycol) (PEG) chain, to the interior of the nanopore, within the large vestibule of the αHL protein pore (Howorka et al. 2000; Movileanu et al. 2000). A biotinyl ligand was permanently attached to the untethered end of the PEG linker. The biotinyl group was able to move across the interior of the nanopore from one side of the membrane to the other, a distance of ~100 Å. If a low-affinity streptavidin mutant was added to the chamber, a distinctive alteration in the single-channel current occurred (Movileanu et al. 2000). This experimental design demonstrated that the single-molecule stochastic sensing of proteins can be fundamentally expanded into the aqueous phase. Inspired by this experimental design, Kong and Muthukumar (2005) performed Langevin molecular dynamics simulations and Poisson-Nernst-Planck calculations to obtain the current fluctuations produced by a single PEG covalently anchored within the large vestibule of the αHL protein pore (Kong and Muthukumar 2005). This computational work revealed the dynamics of the single-molecule captures of the binding protein in the aqueous phase.

A similar approach was employed to examine the binding kinetics of a lectin protein to one or more ligands (Howorka et al. 2004). These experiments showed the ability to detect monovalent and divalent binding events directly. Moreover, a genetically encoded stochastic sensor element has been used for sensing proteins at the single-molecule level (Cheley et al. 2006). In this investigation, an ezymatic peptide inhibitor was engineered in one subunit of the αHL protein nanopore near the pore entrance. Individual binding events of the peptide inhibitor to the catalytic subunit of the cAMP-dependent protein kinase (PKA) were detected through the modulation of the single-channel electrical current. This methodology provided detailed kinetic and thermodynamic information of the peptide–enzyme interactions. This aproach holds potential for rapid screening of kinase inhibitors.

Guan and colleagues have employed an engineered αHL protein pore that contained an iris of seven phenylalanine residues to study the protease activity at the single-molecule level. The hydrophobic trap within the interior of the nanopore functioned as a binding site for hydrophobic residue-containing amyloid-β peptide (Zhao et al. 2009a). Therefore, long-lived current blockades were recorded when this peptide was added to the chamber. Remarkably, the frequency of amyloid-β peptide-induced current blockade was significantly reduced in the presence of trypsin, a serine protease that cleaves peptide bonds after arginine and lysine. Moreover, the frequency of events was also dependent on time, suggesting that this experimental design might be used to obtain a quantitative chemical kinetic information on the enzymatic process. Again, these experiments have proven that the engineered αHL protein pore might be used to examine the enzyme-protein interactions.

4. Detection of proteins with synthetic nanopores

Single-molecule detection of proteins with protein nanopores proves to have a number

of disadvantages and limitations, such as the inability to tune the diameter of the nanopore and limited stability of the protein nanopore-lipid bilayer system under harsh experimental conditions of salt concentration, pH, applied transmembrane potential and other osmotic stress. This daunting challenges might be overcome by the fabrication of solid-state nanopores, which exhibit obvious advantages, such as the accurate control over the inner diameter and the length of the nanopore, as well as a significantly enhanced stability for long periods (Li et al. 2001; Storm et al. 2003; Dekker 2007). Therefore, a recent natural expansion of the single-molecule studies of polypeptides with protein nanopores was the initiation of experiments with synthetic nanopores (Siwy et al. 2005; Han et al. 2006, 2008; Fologea et al. 2007; Sexton et al. 2007; Talaga and Li 2009). In 2005, the first protein biosensor, based on a synthetic nanopore, was designed by Charles Martin's group (Siwy et al. 2005). In this case, the nanopore was a conically shaped gold nanotube reconstituted into a 12-µm-thick poly-ethylene terephthalate membrane. The strategy was to functionalize the small diameter of the conical nanopore with a reactive group. The polymeric membrane was coated with a single layer of Au that functioned as a chemically reactive group for the permanent attachment of a molecular-recognition agent (MRA). Remarkably, such a detector was versatile to a broad range of MRA applications, including protein–ligand (e. g. streptavidin–biotin) and antibody-binding protein (e. g. immunoglobulin IgG–protein-G) complexes. A permanent current blockade was recorded upon the binding of each analyte with the corresponding MRA. The same research team expanded their ability to probe selectively protein analytes with synthetic nanopores (Sexton et al. 2007). In this case, the nanopore element was a PEG-functionalized gold conical nanotube. Bovine serum albumin (BSA) and a Fab fragment from a BSA-polyclonal antibody were detected from the distinctive signature of the current blockades made by the free BSA and the BSA–Fab complex. The Fab fragment was used as an anti-BSA interacting molecule with high binding affinity for the BSA protein. The sensing element was also tested for proteins that do not bind to Fab. It would be interesting to design an experimental protocol that employs proteins with varying binding affinity, so they produce distinctive transient current blockades.

Several research teams pursued single-molecule detection of proteins with single nanopores based on silicon nitride membranes (Han et al. 2006, 2008; Fologea et al. 2007; Talaga and Li 2009). The use of synthetic nanopores has an array of advantageous features, including a greater robustness of the solid-state membranes, the ability to tune the diameter of the nanopore, and the potential for either integration into a *lab-on-a-chip* platform or parallelization for high-throughput devices. Therefore, it is conceivable that the single-molecule studies performed with protein nanopores might be expanded to incorporate a much wider spectrum of applications and harsh conditions of experimentation. For example, temperature-dependence experiments with synthetic nanopores would enable the measurement of kinetic and thermodynamic enthalpies and entropies of the complex formation between a protein and a ligand, revealing information about which process (e. g. enthalpic or entropic) in the ligand-binding protein interactions is dominant. On the other hand, the use of solid-state nanopores has limitations (Griffiths 2008). One obvious challenge is to attach a single functional reactive group within a strategic position of the interior of the synthetic nanopore. This approach would allow transient bindings of a protein analyte to a ligand engineered within a synthetic nanopore.

5. Exploring protein folding and stability

Protein pores and solid-state nanopores were also used as sensor elements for folding and stability features of single proteins. We engineered an exogenous peptide loop within the large vestibule of the aHL protein pore, a cavity with a volume of ~39 500 Å^3 (Jung et al. 2006). The exogenous peptide was an elastine-like-polypeptide (ELP) that undergoes an inverse temperature transition, and with the amino acid sequence $(VPGGG)_5$. It is known that ELPs dehydrate and hydrophobically collapse at greater temperatures, above the inverse transition temperature, and unfold at temperatures below the inverse transition temperature (Urry 1997; Manno et al. 2001; Meyer and Chilkoti 2002, 2004; Urry et al. 2002; Hyun et al. 2004; Zhang et al. 2006). A comprehensive model that shows the conformational fluctuations of the engineered ELP within the large vestibule of the αHL protein pore is depicted in Fig. 6. We probed the hydrophobic collapse of the engineered ELP loop via significant alterations in the temperature-dependent amplitude of the ELP-induced current blockades. For example, at temperatures below its transition temperature, the ELP loop was fully expanded, so it blocked the pore completely, but reversibly (Fig. 7A) (Jung et al. 2006). In contrast, at temperatures above the transition temperature, the ELP excursions into the narrowest region of the interior of the nanopore did not produce a full blockade (Fig. 7B, C). In this case, a substantial flow of ions occurred, because of the hydrophobic collapse of the ELP. These charateristics of the transient ELP-induced

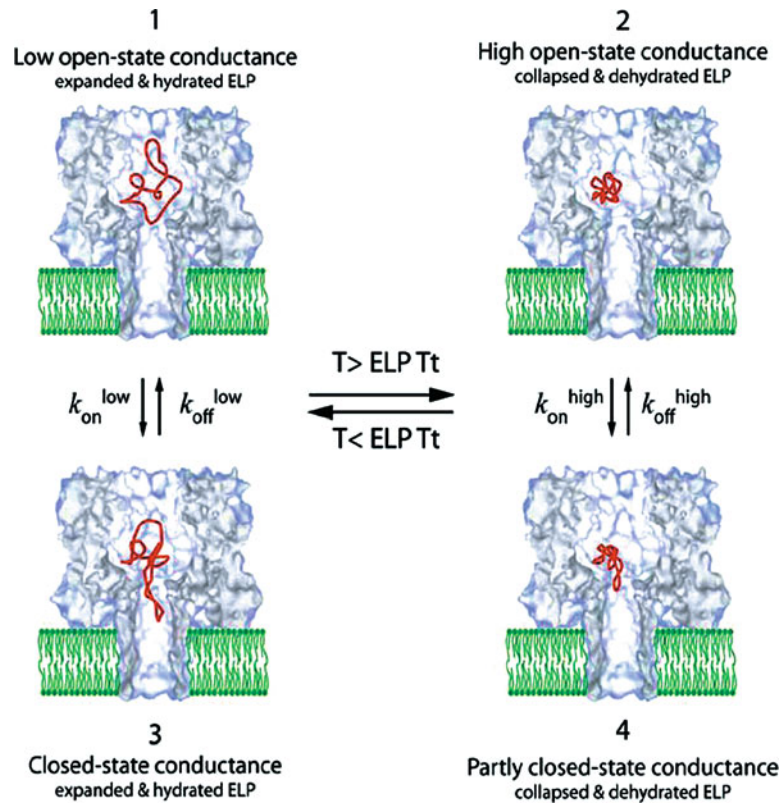


Fig. 6 Model for the temperature-dependent conformational fluctuations of a single ELE loop-containing αHL protein pore. Reproduced, with permission, from Jung et al. (2006)

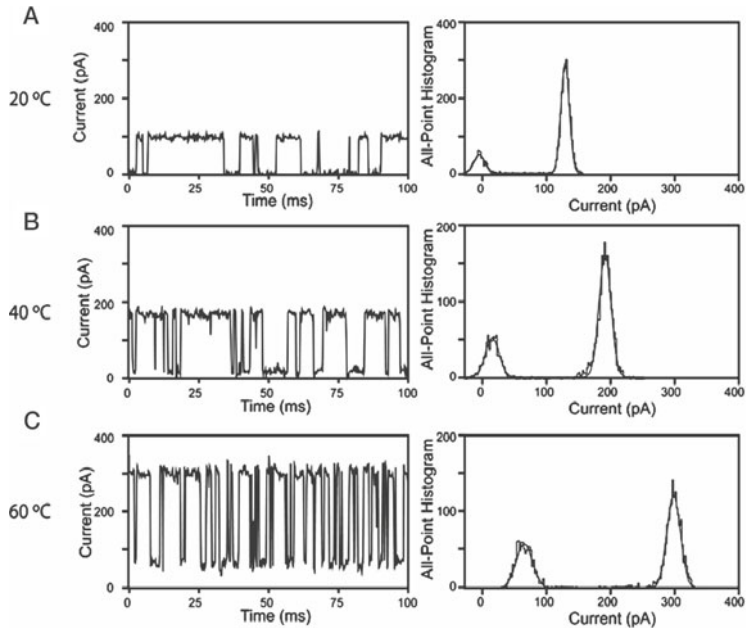


Fig. 7 Single-channel electrical recordings with a single ELP-containing αHL protein pore recorded at different temperatures. Temperature-dependence of the single-channel current through the temperature-responsive pore $E_{10}W_6$, an αHL protein containing a 50-residue long ELP loop, is recorded at (**A**) 20, (**B**) 40 and (**C**) 60 °C. Reproduced, with permission, from Jung et al. (2006)

current blockades were also dependent on the ELP sequence. A test peptide loop with the same length (~50 residues), which was rich in glycine and serine residues, produced a different single-channel electrical signature comprised of highly frequent and very short-lived current spikes. We interpreted that the current spikes were consistent with a substantially increased peptide flexibility because of the glycine residues that have less bulky side chains. These experiments demonstrate our ability to probe conformational fluctuations and flexibilities of the polypeptides engineered within the large vestibule of the nanopore interior. This approach might be expanded to other applications. For example, other engineered αHL pores with responsive polypeptides might be developed. Protein pores with stimulus-activated (e. g., photo-responsive peptides) gating mechanisms might be useful both for basic science studies and the design of new engineered nanopore-based biosensor platforms.

Loic Auvray and colleagues used the αHL protein pore to explore the folding properties of proteins (Oukhaled et al. 2007). These studies were pursued under naturing and denaturing conditions. The protein unfolding was induced by elevated concentrations of guanidinium hydrochloride (Gdm-HCl), a chemical denaturant. The protein under investigation was the 370-residue maltose binding protein of *Escherichia coli* (MBP). The single-channel electrical recordings were performed at different concentrations of Gdm-HCl. In the absence of the denaturing agent, no current blockades were recorded, indicating a large energetic penalty of the protein to interact with the αHL pore. Interestingly, they found protein-induced current blockades in the presence of increased concentrations of Gdm-HCl, suggesting that the MBP protein unfolds at elevated concentrations of chemical denaturant in the chamber. Moreover, the frequency and duration of the protein-induced current blockades was dependent on the concentration of Gdm-HCl in the chamber. This finding was consistent with a clear correlation between the partioning of the MBP protein into the nanopore interior and the folding state of the MBP protein in the aqueous phase.

The thermodynamic stability of a single protein might be altered by a chemical denaturant, temperature or pH. Alternatively, thermodynamic unfolding induced by a chemical denaturant or temperature might be replaced by spontaneous unfolding via single-site mutagenesis within a critical site of the protein. Jeremy Lee and colleagues used the αHL protein pore to demonstrate that the translocation of a protein analyte, a histidine-containing protein (HPr), is strongly dependent upon single-site mutations that modify its folding state (Stefureac et al. 2008). HPr exhibits a low activation free-energy of unfolding HPr (~5 kcal/mol) and this is consistent with the fragile folding state of HPr, as illuminated by single-channel electrical recordings (Stefureac et al. 2008).

Very recently, Talaga and Li (2009) employed silicon nitride-based nanopores to discriminate among folded, partially unfolded and unfolded single proteins (Talaga and Li 2009). The proteins under investigation were β-lactoglobulin and histidine-containing phosphocarrier protein. Interestingly, their findings demonstrated that the proteins translocated through silicon nitride-based nanopores under linear or looped conformation and the applied electrical force can contribute to the protein unfolding during transit.

6. Probing protein-nucleic acid complexes with nanopores

In the previous sections of this chapter, we showed that nanopore research has been employed in the examination of recognition templates, such as small ligand–protein (Movileanu et al. 2000; Howorka et al. 2004; Mayer et al. 2008), peptide–protein (Xie et al. 2005; Cheley et al. 2006) and antibody–protein (Siwy et al. 2005; Sexton et al. 2007; Han et al. 2008) complexes. However, several groups have started to investigate the subtle interactions of the binding proteins with their nucleic acid targets (Astier et al. 2007; Benner et al. 2007; Hornblower et al. 2007; Zhao et al. 2007; Smeets et al. 2009). Nucleic acid-protein interaction is fundamental and ubiquitous in modern biology and medical biotechnology. Therefore, with more adaptation of the nanopore probe techniques, parallel nanopore-based arrays might be used in the future for high-throughput devices in protein sensing and proteomics. Li-Qun Gu's group used nanopores to investigate aptamers and aptamer-binding protein complexes (Shim and Gu 2008; Shim et al. 2009). Aptamers are short nucleic acid sequences that bind with enhanced binding affinity to small organic molecules, polypeptides, and proteins. Li-Qun Gu and colleagues examined the single-molecule detection of thrombin-binding aptamer (TBA), a G-quadruplex nucleic acid, in the presence of divalent cations (Shim and Gu 2008). In this case, the large vestibule of the pore was used to trap single thrombin molecules through the electrophoretic insertion of a single G-quadruplex oligonucleotide. This study demonstrates a remarkable power of the nanopore detector for sensing the DNA/RNA aptamers. There are numerous advantages of the DNA/RNA aptamers as molecular recognition elements in biosensing when compared with traditional detection methodologies using enzymes and antibodies (Song et al. 2008). These include the following: i) enhanced specificity, selectivity, and affinity from small molecules to folded protein domains, enabling the design of high-performance aptamer-based biosensors; ii) extraordinary chemical stability, thereby offering exceptional flexibility and convenience in their design; iii) conformational changes upon target binding, conferring a great flexibility in designing novel

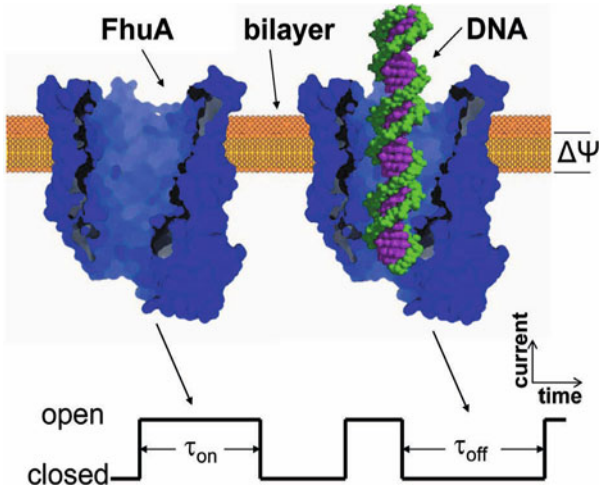


Fig. 8 Translocation of double-stranded DNA (dsDNA) through a single cork-free FhuA-based nanopore placed in a planar lipid bilayer. dsDNA is electrically driven by the transmembrane potential ($\Delta\psi$).The frequency of occurrence (~$1/\tau_{on}$) of the events reveals the concentration of dsDNA in the bulk aqueous phase, whereas the current signature (the mean duration τ_{off} and amplitude of the transient current blockades) reveals its identity (length and structure, respectively). Stochastic sensing of dsDNA provides useful kinetic data that are difficult to obtain with other traditional techniques. In a simple equilibrium, $\tau_{off} = 1/k_{off}$, where k_{off} is the dissociation rate constant, and $\tau_{on} = 1/k_{on}[C_{DNA}]$, where k_{on} is the association rate constant and $[C_{DNA}]$ is the DNA concentration in the aqueous phase

biosensors with unprecedented detection sensitivity.

Other groups used protein and solid-state nanopores for quantitative explorations of the biophysical and biochemical characteristics of the complexes between proteins and nucleic acids, such as the voltage threshold for complex dissociation (Zhao et al. 2007), the sequence-specific sensitivity detection of individual DNA polymerase (Benner et al. 2007), the underlying equilibrium kinetics and thermodynamics of protein–nucleic acid interactions (Hornblower et al. 2007), the motor activity of RNA-processing enzymes (Astier et al. 2007), the transit time of the protein–nucleic acid complex within the nanopore (Smeets et al. 2009), and DNA polymerase activity (Cockroft et al. 2008). In general, these complexes are much bulkier than single-stranded DNA (ssDNA) or unfolded polypeptides. Therefore, nanopores designed for these applications should have

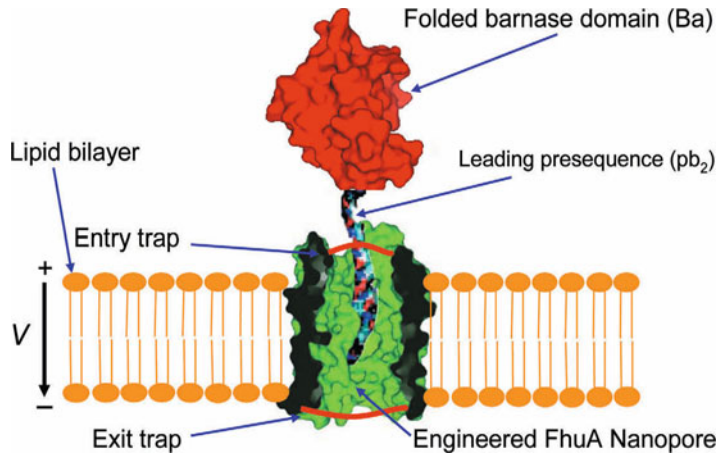


Fig. 9 Clamping a single pb_2-Ba protein with electrostatic traps. The two electrostatic traps (in red), made from aspartic acids, are shown on the entry and exit of the cork-free FhuA-based nanopore. Partitioning of a single pb_2-Ba protein into the nanopore lumen should be observed by a transient current blockade. Both proteins are represented at the same scale

a larger internal diameter, in the range of several nanometers. For example, ferric hydroxamate uptake component A (FhuA) (Ferguson et al. 1998; Locher et al. 1998), a monomeric β-barrel protein might be used for stochastic sensing of double-stranded DNA (dsDNA) (Fig. 8) and proteins (Fig. 9). FhuA is a member of the superfamily of bacterial outer membrane proteins. The principles underlying the functionality of this sensor element are inspired by natural processes: translocation of phage DNA (Letellier et al. 1999; Gurnev et al. 2006) and folded proteins (Hill et al. 1998; Hinnah et al. 2002; Muro et al. 2003; Becker et al. 2005) through a β-barrel protein channel.

Concluding remarks

In this chapter, we discussed a number of ways by which short polypeptides and small folded proteins might be examined using single-molecule nanopore probe techniques and stochastic sensing. The underlying detection principle can be applied to both natural and synthetic nanopores. The protein pores have the advantageous feature of site-directed functionalization and chemical modification within strategic positions of the nanopore interior. In analogy to their biological counterparts, synthetic nanopores are used as stochastic sensors to probe single proteins and their complexes with interacting ligands (Han et al. 2006, 2008; Talaga and Li 2009). The detection mechanisms of polypeptides and folded proteins include genetic engineering and chemical modification of protein nanopores, functionalization of the inner surface of solid-state nanopores and electrophoretic capture through movable nucleic acid arms. Critically important, nanopore probe techniques enable exploration of specific features of proteins such as their enzymatic activity (Xie et al. 2005; Cheley et al. 2006) and binding affinity (Movileanu et al. 2000; Howorka et al. 2004) at the single-molecule level. It is worth mentioning that this methodology enables the determination of these features without resorting to labeling of the interacting molecules (Movileanu 2009), avoiding several disadvantages of the detection using fluorescent dyes.

On the other hand, coupling a nanopore probe technique to an optical or fluorescence platform is expected to enhance our knowledge about the underlying mechanisms of protein translocation. The translocation of proteins through a single protein nanopore represents a critical process in protein unfolding through proteosome channel and protein traffic across mitochondrial cell membranes (Matouschek and Glick 2001; Huang et al. 2002; Matouschek 2003; Prakash and Matouschek 2004; Shariff et al. 2004; Wilcox et al. 2005; Mohammad et al. 2008; Inobe et al. 2008; Prakash et al. 2009). However, in these processes the translocation and unfolding of the protein are two intimately related mechanisms. We anticipate that a biophysical instrument that consists of a solid-state nanopore and an optical trap (Keyser et al. 2006; Dekker 2007) has the potential to illuminate the balance of driving and unfolding forces that act on a protein during its translocation in a linear manner through a nanopore. Such efforts will certainly lead to a better understanding of how proteins undergo conformational transitions when they traverse a nanopore. The experimental studies devoted to protein translocations will inspire new theoretical and computational biophysics studies on protein translocation through nanopores (Li and Makarov 2003, 2004a,b; Kirmizialtin et al. 2004, 2005, 2006; Huang et al. 2005; Tian and Andricioaei 2005; Goodrich et al. 2007; Huang and Makarov 2008a,b; Li et al. 2006; Makarov 2007, 2008).

Moreover, nanopore probe techniques are successfully used to examine the folding state and the flexibility of single proteins (Goodrich et al. 2007; Oukhaled et al. 2007). Despite these recent advances in using nanopores for the exploration of single proteins, more experimentation is needed to better understand the subtle interactions of single proteins with the interior of the nanopores. For example, it is not clearly understood how the flexibility and stability of the protein is altered upon its partitioning into the nanopore interior. How do the proteins interact with the slightly hydrophobic surface of the silicon nitride-based nanopores? What is the magnitude of the electrostatic force between polypeptides and the surface of the protein nanopores? Do the proteins traverse a nanopore in unfolded or partly unstructured conformation? What is the time scale of the

protein unfolding in the interior of the nanopore? It is likely that more experimental work and future advances of nanopore technology will be employed to tackle numerous questions that are pertinent to protein translocation.

Acknowledgements

I am grateful to researchers in my group, Carl P. Goodrich, Aaron J. Wolfe, Mohammad M. Mohammad, David J. Niedzwiecki, Khalil R. Howard, and Robert Bikwemu for their work performed at Syracuse University. I also thank my colleagues Dmitrii Makarov (The University of Texas at Austin, Austin, TX), Andreas Matouschek (Northwestern University, Evanston, IL) and Anatoly Kolomeisky (Rice University, Houston, TX) for stimulating discussions. This work was funded in part by grants from the US National Science Foundation (DMR-0706517 and DMR-1006332), the National Institutes of Health (GM088403-01) and by the Syracuse Biomaterials Institute (SBI).

References

Akeson M, Branton D, Kasianowicz JJ, Brandin E, Deamer DW (1999) Microsecond time-scale discrimination among polycytidylic acid, polyadenylic acid, and polyuridylic acid as homopolymers or as segments within single RNA molecules. Biophys J 77: 3227–3233

Astier Y, Kainov DE, Bayley H, Tuma R, Howorka S (2007) Stochastic detection of motor protein-RNA complexes by single-channel current recording. Chemphyschem 8: 2189–2194

Bayley H (1999) Designed membrane channels and pores. Curr Opin Biotechnol 10: 94–103

Bayley H (2006) Sequencing single molecules of DNA. Curr Opin Chem Biol 10: 628–637

Bayley H, Braha O, Cheley S, Gu LQ (2004) Engineered nanopores. In: Niemeyer CMaMCA (ed) NanoBiotechnology, pp 93–112. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Bayley H, Braha O, Gu LQ (2000) Stochastic sensing with protein pores. Adv Mater 12: 139–142

Bayley H, Cremer PS (2001) Stochastic sensors inspired by biology. Nature 413: 226–230

Bayley H, Jayasinghe L (2004) Functional engineered channels and pores (Review). Mol Membr Biol 21: 209–220

Bayley H, Martin CR (2000) Resistive-pulse sensing – From microbes to molecules. Chem Rev 100: 2575–2594

Becker L, Bannwarth M, Meisinger C, Hill K, Model K, Krimmer T, Casadio R, Truscott KN, Schulz GE, Pfanner N, Wagner R (2005) Preprotein translocase of the outer mitochondrial membrane: reconstituted tom40 forms a characteristic TOM Pore. J Mol Biol 353: 1011–1020

Benner S, Chen RJ, Wilson NA, Abu-Shumays R, Hurt N, Lieberman KR, Deamer DW, Dunbar WB, Akeson M (2007) Sequence-specific detection of individual DNA polymerase complexes in real time using a nanopore. Nat Nanotechnol 2: 718–724

Bezrukov SM (2000) Ion channels as molecular Coulter counters to probe metabolite transport. J Membr Biol 174: 1–13

Bezrukov SM, Kasianowicz JJ (1993) Current noise reveals protonation kinetics and number of ionizable sites in an open protein ion channel. Phys Rev Lett 70: 2352–2355

Bezrukov SM, Vodyanoy I, Brutyan RA, Kasianowicz JJ (1996) Dynamics and free energy of polymers partitioning into a nanoscale pore. Macromolecules 29: 8517–8522

Branton D, Deamer DW, Marziali A, Bayley H, Benner SA, Butler T, Di Ventra M, Garaj S, Hibbs A, Huang X, Jovanovich SB, Krstic PS, Lindsay S, Ling XS, Mastrangelo CH, Meller A, Oliver JS, Pershin YV, Ramsey JM, Riehn R, Soni GV, Tabard-Cossa V, Wanunu M, Wiggin M Schloss JA (2008) The potential and challenges of nanopore sequencing. Nat Biotechnol 26: 1146–1153

Butler TZ, Gundlach JH, Troll M (2007) Ionic current blockades from DNA and RNA molecules in the alpha-hemolysin nanopore. Biophys J 93: 3229–3240

Butler TZ, Pavlenok M, Derrington I, Niederweis M, Gundlach JH (2008) Single-molecule DNA detection with an engineered MspA protein nanopore. Proc Natl Acad Sci U S A 105: 20647–20652

Cheley S, Xie H, Bayley H (2006) A genetically encoded pore for the stochastic detection of a protein kinase. Chembiochem 7: 1923–1927

Cockroft SL, Chu J, Amorin M, Ghadiri MR (2008) A single-molecule nanopore device detects DNA polymerase activity with single-nucleotide resolution. J Am Chem Soc 130: 818–820

Comer J, Dimitrov V, Zhao Q, Timp G, Aksimentiev A (2009) Microscopic mechanics of hairpin DNA translocation through synthetic nanopores. Biophys J 96: 593–608

Deamer DW, Akeson M (2000) Nanopores and nucleic acids: prospects for ultrarapid sequencing. Trends Biotechnol 18: 147–151

Deamer DW, Branton D (2002) Characterization of nucleic acids by nanopore analysis. Acc Chem Res 35: 817–825

Dekker C (2007) Solid-state nanopores. Nature Nanotechnology 2: 209–215

Dorvel B, Sigalov G, Zhao Q, Comer J, Dimitrov V, Mirsaidov U, Aksimentiev A, Timp G (2009) Analyzing the forces binding a restriction endonuclease to DNA using a synthetic nanopore. Nucleic Acids Res 37: 4170–4179

Ferguson AD, Hofmann E, Coulton JW, Diederichs K, Welte W (1998) Siderophore-mediated iron transport: crystal structure of FhuA with bound lipopolysaccharide. Science 282: 2215–2220

Fologea D, Ledden B, McNabb DS, Li J (2007) Electrical characterization of protein molecules by a solid-state nanopore. Appl Phys Lett 91: 053901

Goodrich CP, Kirmizialtin S, Huyghues-Despointes BM, Zhu AP, Scholtz, J. M., Makarov DE, Movileanu L (2007) Single-molecule electrophoresis of beta-hairpin peptides by electrical recordings and Langevin dynamics simulations. J Phys Chem B 111: 3332–3335

Griffiths J (2008) The realm of the nanopore: interest in nanoscale research has skyrocketed, and the humble pore has become a king. Anal Chem 80: 23–27

Gurnev PA, Oppenheim AB, Winterhalter M, Bezrukov SM (2006) Docking of a single phage lambda to its membrane receptor maltoporin as a time-resolved event. J Mol Biol 359: 1447–1455

Han A, Creus M, Schurmann G, Linder V, Ward TR, de Rooij NF, Staufer U (2008) Label-free detection of single protein molecules and protein-protein interactions using synthetic nanopores. Anal Chem 80: 4651–4658

Han A, Schurmann G, Monding G, Bitterli RA, de Rooij NF, Staufer U (2006) Sensing protein molecules using nanofabricated pores. Appl Phys Lett 88: 093901

Hayden T (2008) Threading proteins through a nanopore needle. Anal Chem 80: 3955

Henrickson SE, Misakian M, Robertson B, Kasianowicz JJ (2000) Driven DNA transport into an asymmetric nanometer-scale pore. Phys Rev Lett 85: 3057–3060

Hill K, Model K, Ryan MT, Dietmeier K, Martin F, Wagner R, Pfanner N (1998) Tom40 forms the hydrophilic channel of the mitochondrial import pore for preproteins. Nature 395: 516–521

Hinnah SC, Wagner R, Sveshnikova N, Harrer R, Soll J (2002) The chloroplast protein import channel Toc75: Pore properties and interaction with transit peptides. Biophys J 83: 899–911

Hornblower B, Coombs A, Whitaker RD, Kolomeisky A, Picone SJ, Meller A, Akeson M (2007) Single-molecule analysis of DNA-protein complexes using nanopores. Nat Methods 4: 315–317

Howorka S, Bayley H (2002) Probing distance and electrical potential within a protein pore with tethered DNA. Biophys J 83: 3202–3210

Howorka S, Cheley S, Bayley H (2001a) Sequence-specific detection of individual DNA strands using engineered nanopores. Nat Biotechnol 19: 636–639

Howorka S, Movileanu L, Braha O, Bayley H (2001b) Kinetics of duplex formation for individual DNA strands within a single protein nanopore. Proc Natl Acad Sci U S A 98: 12996–13001

Howorka S, Movileanu L, Lu XF, Magnon M, Cheley S, Braha O, Bayley H (2000) A protein pore with a single polymer chain tethered within the lumen. J Am Chem Soc 122: 2411–2416

Howorka S, Nam J, Bayley H, Kahne D (2004) Stochastic detection of monovalent and bivalent protein-ligand interactions. Angew Chem Int Ed Engl 43: 842–846

Howorka S, Siwy Z (2008) Nanopores: generation, engineering and single-molecule applications. In: Hinterdorfer P (ed) Handbook of single-molecule biophysics, pp In press. Springer-Verlag, Heidelberg, New York

Howorka S, Siwy Z (2009) Nanopore analytics: sensing of single molecules. Chem Soc Rev 38: 2360–2384

Huang L, Kirmizialtin S, Makarov DE (2005) Computer simulations of the translocation and unfolding of a protein pulled mechanically through a pore. J Chem Phys 123: 124903

Huang L, Makarov DE (2008a) The rate constant of polymer reversal inside a pore. J Chem Phys 128: 114903

Huang L, Makarov DE (2008b) Translocation of a knotted polypeptide through a pore. J Chem Phys 129: 121107

Huang S, Ratliff KS, Matouschek A (2002) Protein unfolding by the mitochondrial membrane potential. Nat Struct Biol 9: 301–307

Hyun J, Lee WK, Nath N, Chilkoti A, Zauscher S (2004) Capture and release of proteins on the

nanoscale by stimuli-responsive elastin-like polypeptide "switches". J Am Chem Soc 126: 7330–7335

Inobe T, Kraut DA, Matouschek A (2008) How to pick a protein and pull at it. Nat Struct Mol Biol 15: 1135–1136

Jung Y, Bayley H, Movileanu L (2006) Temperature-responsive protein pores. J Am Chem Soc 128: 15332–15340

Kang XF, Gu LQ, Cheley S, Bayley H (2005) Single protein pores containing molecular adapters at high temperatures. Angew Chem Int Ed Engl 44: 1495–1499

Kasianowicz JJ (2004) Nanopores: flossing with DNA. Nat Mater 3: 355–356

Kasianowicz JJ, Bezrukov SM (1995) Protonation dynamics of the alpha-toxin ion-channel from spectral-analysis of ph-dependent current fluctuations. Biophys J 69: 94–105

Kasianowicz JJ, Brandin E, Branton D, Deamer DW (1996) Characterization of individual polynucleotide molecules using a membrane channel. Proc Natl Acad Sci U S A 93: 13770–13773

Keyser UF, Koeleman BN, van Dorp S, Krapf D, Smeets RMM, Lemay SG, Dekker NH, Dekker C (2006) Direct force measurements on DNA in a solid-state nanopore. Nature Physics 2: 473–477

Kim M-J, Wanunu M, Bell CD, Meller A (2006) Rapid fabrication of uniform size nanopores and nanopore arrays for parallel DNA analysis. Adv Mater 18: 3149–3153

Kirmizialtin S, Ganesan V, Makarov DE (2004) Translocation of a beta-hairpin-forming peptide through a cylindrical tunnel. J Chem Phys 121: 10268–10277

Kirmizialtin S, Huang L, Makarov DE (2005) Topography of the free-energy landscape probed via mechanical unfolding of proteins. J Chem Phys 122: 234915

Kirmizialtin S, Huang L, Makarov DE (2006) Computer simulations of protein translocation. Phys Stat Sol (b) 243: 2038–2047

Kong CY, Muthukumar M (2005) Simulations of stochastic sensing of proteins. J Am Chem Soc 127: 18252–18261

Krasilnikov OV, Bezrukov SM (2004) Polymer partitioning from nonideal solutions into protein voids. Macromolecules 37: 2650–2657

Krasilnikov OV, Rodrigues CG, Bezrukov SM (2006) Single polymer molecules in a protein nanopore in the limit of a strong polymer-pore attraction. Phys Rev Lett 97: 018301

Letellier L, Plancon L, Bonhivers M, Boulanger P (1999) Phage DNA transport across membranes. Res Microbiol 150: 499–505

Li J, Stein D, McMullan C, Branton D, Aziz MJ, Golovchenko JA (2001) Ion-beam sculpting at nanometre length scales. Nature 412: 166–169

Li PC, Huang L, Makarov DE (2006) Mechanical unfolding of segment-swapped protein G dimer: results from replica exchange molecular dynamics simulations. J Phys Chem B Condens Matter Mater Surf Interfaces Biophys 110: 14469–14474

Li PC, Makarov DE (2003) Theoretical studies of the mechanical unfolding of the muscle protein titin: Bridging the time-scale gap between simulation and experiment. J Chem Phys 119: 9260–9268

Li PC, Makarov DE (2004a) Simulation of the mechanical unfolding of ubiquitin: probing different unfolding reaction coordinates by changing the pulling geometry. J Chem Phys 121: 4826–4832

Li PC, Makarov DE (2004b) Ubiquitin-like protein domains show high resistance to mechanical unfolding similar to that of the 127 domain in titin: Evidence from simulations. J Phys Chem B 108: 745–749

Locher KP, Rees B, Koebnik R, Mitschler A, Moulinier L, Rosenbusch JP, Moras D (1998) Transmembrane signaling across the ligand-gated FhuA receptor: crystal structures of free and ferrichrome-bound states reveal allosteric changes. Cell 95: 771–778

Makarov DE (2007) Unraveling individual molecules by mechanical forces: theory meets experiment. Biophys J 92: 4135–4136

Makarov DE (2008) Computer simulations and theory of protein translocation. Acc Chem Res 42: 281–289

Manno M, Emanuele A, Martorana V, San Biagio PL, Bulone D, Palma-Vittorelli MB, McPherson DT, Xu J, Parker TM, Urry DW (2001) Interaction of processes on different length scales in a bioelastomer capable of performing energy conversion. Biopolymers 59: 51–64

Martin CR, Siwy ZS (2007) Chemistry. Learning nature's way: biosensing with synthetic nanopores. Science 317: 331–332

Mathe J, Aksimentiev A, Nelson DR, Schulten K, Meller A (2005) Orientation discrimination of single-stranded DNA inside the alpha-hemolysin membrane channel. Proc Natl Acad Sci U S A 102: 12377–12382

Matouschek A (2003) Protein unfolding – an im-

portant process in vivo? Curr Opin Struct Biol 13: 98–109

Matouschek A, Glick BS (2001) Barreling through the outer membrane. Nat Struct Biol 8: 284–286

Mayer M, Semetey V, Gitlin I, Yang J, Whitesides GM (2008) Using ion channel-forming peptides to quantify protein-ligand interactions. J Am Chem Soc 130: 1453–1465

McNally B, Wanunu M, Meller A (2008) Electromechanical unzipping of individual DNA molecules using synthetic sub-2 nm pores. Nano Lett 8: 3418–3422

Meller A (2003) Dynamics of polynucleotide transport through nanometre-scale pores. J Phys Condens Matter 15: R581-R607

Meller A, Nivon L, Brandin E, Golovchenko J, Branton D (2000) Rapid nanopore discrimination between single polynucleotide molecules. Proc Natl Acad Sci U S A 97: 1079–1084

Meyer DE, Chilkoti A (2002) Genetically encoded synthesis of protein-based polymers with precisely specified molecular weight and sequence by recursive directional ligation: Examples from the elastin-like polypeptide system. Biomacromolecules 3: 357–367

Meyer DE, Chilkoti A (2004) Quantification of the effects of chain length and concentration on the thermal behavior of elastin-like polypeptides. Biomacromolecules 5: 846–851

Mohammad MM, Movileanu L (2008) Excursion of a single polypeptide into a protein pore: simple physics, but complicated biology. Eur Biophys J 37: 913–925

Mohammad MM, Prakash S, Matouschek A, Movileanu L (2008) Controlling a single protein in a nanopore through electrostatic traps. J Am Chem Soc 130: 4081–4088

Movileanu L (2008) Squeezing a single polypeptide through a nanopore. Soft Matter 4: 925–931

Movileanu L (2009) Interrogating single proteins through nanopores: challenges and opportunities. Trends Biotechnol 27: 333–341

Movileanu L, Cheley S, Howorka S, Braha O, Bayley H (2001) Location of a constriction in the lumen of a transmembrane pore by targeted covalent attachment of polymer molecules. J Gen Physiol 117: 239–251

Movileanu L, Howorka S, Braha O, Bayley H (2000) Detecting protein analytes that modulate transmembrane movement of a polymer chain within a single protein pore. Nat Biotechnol 18: 1091–1095

Movileanu L, Schmittschmitt JP, Scholtz JM, Bayley H (2005) Interactions of the peptides with a protein pore. Biophys J 89: 1030–1045

Muro C, Grigoriev SM, Pietkiewicz D, Kinnally KW, Campo ML (2003) Comparison of the TIM and TOM channel activities of the mitochondrial protein import complexes. Biophys J 84: 2981–2989

Oukhaled G, Mathe J, Biance A-L, Bacri L, Betton J-M, Lairez D, Pelta J, Auvray L (2007) Unfolding of proteins and long transient conformations detected by single nanopore recording. Phys Rev Lett 98: 158101

Prakash S, Inobe T, Hatch AJ, Matouschek A (2009) Substrate selection by the proteasome during degradation of protein complexes. Nat Chem Biol 5: 29–36

Prakash S, Matouschek A (2004) Protein unfolding in the cell. Trends Biochem Sci 29: 593–600

Rhee M, Burns MA (2007) Nanopore sequencing technology: nanopore preparations. Trends Biotechnol 25: 174–181

Rodrigues CG, Machado DC, Chevtchenko SF, Krasilnikov OV (2008) Mechanism of KCl enhancement in detection of nonionic polymers by nanopore sensors. Biophys J 95: 5186–5192

Sanchez-Quesada J, Ghadiri MR, Bayley H, Braha O (2000) Cyclic peptides as molecular adapters for a pore-forming protein. J Am Chem Soc 122: 11757–11766

Sexton LT, Horne LP, Sherrill SA, Bishop GW, Baker LA, Martin CR (2007) Resistive-pulse studies of proteins and protein/antibody complexes using a conical nanotube sensor. J Am Chem Soc 129: 13144–13152

Shariff K, Ghosal S, Matouschek A (2004) The force exerted by the membrane potential during protein import into the mitochondrial matrix. Biophys J 86: 3647–3652

Shim JW, Gu LQ (2008) Encapsulating a single g-quadruplex aptamer in a protein nanocavity. J Phys Chem B 112: 8354–8360

Shim JW, Tan Q, Gu LQ (2009) Single-molecule detection of folding and unfolding of the G-quadruplex aptamer in a nanopore nanocavity. Nucleic Acids Res 37: 972–982

Siwy Z, Trofin L, Kohli P, Baker LA, Trautmann C, Martin CR (2005) Protein biosensors based on biofunctionalized conical gold nanotubes. J Am Chem Soc 127: 5000–5001

Smeets RM, Kowalczyk SW, Hall AR, Dekker NH, Dekker C (2009) Translocation of RecA-Coated

Double-Stranded DNA through Solid-State Nanopores. Nano Lett 9: 3089–3096

Song LZ, Hobaugh MR, Shustak C, Cheley S, Bayley H, Gouaux JE (1996) Structure of staphylococcal alpha-hemolysin, a heptameric transmembrane pore. Science 274: 1859–1

Song S, Wang L, Li J, Zhao J, Fan C (2008) Aptamer-based biosensor. Trends Anal Chem 27: 108–117

Stefureac R, Long YT, Kraatz HB, Howard P, Lee JS (2006) Transport of alpha-helical peptides through alpha-hemolysin and aerolysin pores. Biochemistry 45: 9172–9179

Stefureac R, Waldner L, Howard P, Lee JS (2008) Nanopore analysis of a small 86-residue protein. Small 4: 59–63

Stefureac RI, Lee JS (2008) Nanopore analysis of the folding of zinc fingers. Small 4: 1646–1650

Storm AJ, Chen JH, Ling XS, Zandbergen HW, Dekker C (2003) Fabrication of solid-state nanopores with single-nanometre precision. Nat Mater 2: 537–540

Sutherland T. C., Long YT, Stefureac RI, Bediako-Amoa I, Kraatz HB, Lee JS (2005) Structure of peptides investigated by nanopore analysis. Nano Lett 4: 1273–1277

Talaga DS, Li J (2009) Single-molecule protein unfolding in solid state nanopores. J Am Chem Soc 131: 9287–9297

Tian P, Andricioaei I (2005) Repetitive pulling catalyzes co-translocational unfolding of barnase during import through a mitochondrial pore. J Mol Biol 350: 1017–1034

Urry DW (1997) Physical chemistry of biological free energy transduction as demonstrated by elastic protein-based polymers. J Phys Chem B 101: 11007–11028

Urry DW, Hugel T, Seitz M, Gaub HE, Sheiba L, Dea J, Xu J, Parker T (2002) Elastin: a representative ideal protein elastomer. Phil Trans R Soc Lond, B, Biol Sci 357: 169–184

Wanunu M, Sutin J, McNally B, Chow A, Meller A (2008) DNA translocation governed by interactions with solid-state nanopores. Biophys J 95: 4716–4725

Wilcox AJ, Choy J, Bustamante C, Matouschek A (2005) Effect of protein structure on mitochondrial import. Proc Natl Acad Sci U S A 102: 15435–15440

Wolfe AJ, Mohammad MM, Cheley S, Bayley H, Movileanu L (2007) Catalyzing the translocation of polypeptides through attractive interactions. J Am Chem Soc 129: 14034–14041

Xie H, Braha O, Gu LQ, Cheley S, Bayley H (2005) Single-molecule observation of the catalytic subunit of cAMP-dependent protein kinase binding to an inhibitor peptide. Chem Biol 12: 109–120

Zhang Y, Trabbic-Carlson K, Albertorio F, Chilkoti A, Cremer PS (2006) Aqueous two-phase system formation kinetics for elastin-like polypeptides of varying chain length. Biomacromolecules 7: 2192–2199

Zhao Q, de Zoysa RS, Wang D, Jayawardhana DA, Guan X (2009a) Real-time monitoring of peptide cleavage using a nanopore probe. J Am Chem Soc 131: 6324–6325

Zhao Q, Jayawardhana DA, Wang D, Guan X (2009b) Study of peptide transport through engineered protein channels. J Phys Chem B 113: 3572–3578

Zhao Q, Sigalov G, Dimitrov V, Dorvel B, Mirsaidov U, Sligar S, Aksimentiev A, Timp G (2007) Detecting SNPs using a synthetic nanopore. Nano Lett 7: 1680–1685

A numerical approach to surface plasmon resonance sensor design with high sensitivity using single and bimetallic film structures

26

Hyuk Rok Gwon, Seong Hyuk Lee

Contents

Abstract

Up to now, surface plasmon resonance (SPR) has been widely studied for biological and chemical sensing applications. The present study investigates numerically the optical characteristics for the single Au film and bimetallic Ag-Au film SPR configurations by using the multiple beam interference matrix (MBIM) method. The Kretschmann configuration is used for excitation of surface plasmon wave (SPW) with the use of angular modulation. The estimated results of the reflectance, the phase shift, and the magnetic field intensity enhancement are provided for predicting the optimum metal film thickness which can yield highest sensitivity in measuring the SPR signal. As a result, it is observed that the optimum thickness of a single Au film is 52 nm, whereas for a bimetallic Ag-Au film, the optimum film thickness values are 36 nm for Ag and 5 nm for Au, respectively. Moreover, it is concluded that the phase shift would be remarkably appropriate in detecting the SPR sensor signal because of the high sensitivity it offers compared to other values, such as the full width half maximum (FWHM) of reflectance and the enhancement of magnetic field intensity.

Hyuk Rok Gwon
Chung-Ang University
Heuksuk-Dong 221, Dongjak-Gu
Seoul 156-756, Korea
e-mail: hrgwon99@gmail.com

1. Introduction

Surface plasmon resonance (SPR) is an optical phenomenon which because of its high sensitivity has gained enormous importance in bio-sensing and lab-on-a-chip sensors. The effect is based on the excitation of free electrons in very thin layers of metals like silver and gold. The excitation of surface electromagnetic resonant waves is induced by incident light (typically visible or infrared light) and to a large extent determined by the properties of the interface between

the metal layer and a medium with different refractive index, like glass. SPR leads to a decrease of the intensity of reflected light seen as a well defined minimum of reflectance at a specific incident angle. The resonance conditions are changed when the interface between the two materials of different refractive index changes chemically or physically by the binding of analytes or the adsorption of material onto the thin metal film (which serves as the receptive sensor layer and may be coated with an affinity ligand). Such changes are seen as a shift in the resonance angle, the angle of reflection minimum (absorption maximum).

Detecting molecular adsorption, SPR has been an essential technique in chemical- and bio-sensing. According to a demonstration by Homola et al. (1999), the SPR technique can be used for in-situ and label-free measurements with very high sensing sensitivity. The use of SPR for gas detection and bio-sensing was already demonstrated almost 30 years ago by Nylander and Liedberg (1982, 1983, 1995). In resonant mode, the reflectance, the phase shift, and the magnitude of the electro-magnetic (EM) field are drastically changed. Most researchers have measured reflectance in finding the SPR angle (angular modulation SPR system) or the SPR wavelength (wavelength modulation SPR system). In angular modulation SPR sensor system, when a light beam impinges onto a metal film at a specific (resonance) angle, the surface plamons are set to resonate with the light. Therefore, the resonance results in the absorption of light. Likewise, the SPR wavelength means the specific wavelength of incidence light that has minimum reflectance in wavelength modulation SPR system. To date, some uses of phase shift in SPR sensor application have been reported. Ho et al. (2007) showed that with the measurement of phase shift, the sensitivity of a SPR biosensor was substantially enhanced on the basis of multi-pass interferometry. And Lee et al. (2007) used the phase shift for finding the variation of refractive index with the phase quadrature interferometry. In addition, Chen et al. (2007) showed small differences in the refractive indices of solution by the phase shift with the interferometric method. There also exists the possibility to measure the magnetic field on an SPR sensor: by using magneto-resistive magnetometer, hall devices, and magneto-optic imager, one is able to measure magnetic field quantitatively as demonstrated previously, Göktepe (2001), Tian et al. (2005), Joubert and Pinassaud (2006), and Cui et al. (2007).

Since the surface plasmons (SPs) are collective oscillations of free electrons in metallic films, some metal such as gold, silver, and aluminium can be used to generate SPR signals. Even if a surface instability problem exists, the use of a silver film is capable of yielding much higher resolution compared with a gold film in measuring SPR angles. For this reason, many researchers have attempted to utilize bimetallic Ag-Au film structures to achieve better surface stability, lower chemical resistance, and higher sensitivity.

Furthermore, the change in the reflectance, the phase shift, and the magnitude of the electromagnetic field in SPR sensors depends intimately on the film thickness and the optical properties of the metal layers, such as the reflective index n, and the extinction coefficient k. Therefore, it is especially meaningful to study the influence of types of metal and film thickness on the improvement of sensitivity prior to designing SPR sensors. Thus, the present study provides numerical results for two different SPR configurations: one of which involves a single Au film deposited on a glass prism, and the other employs a bimetallic film structure designed with the use of silver and gold films as shown in Fig. 1 which will be explained in detail in section 2. In addition, the method

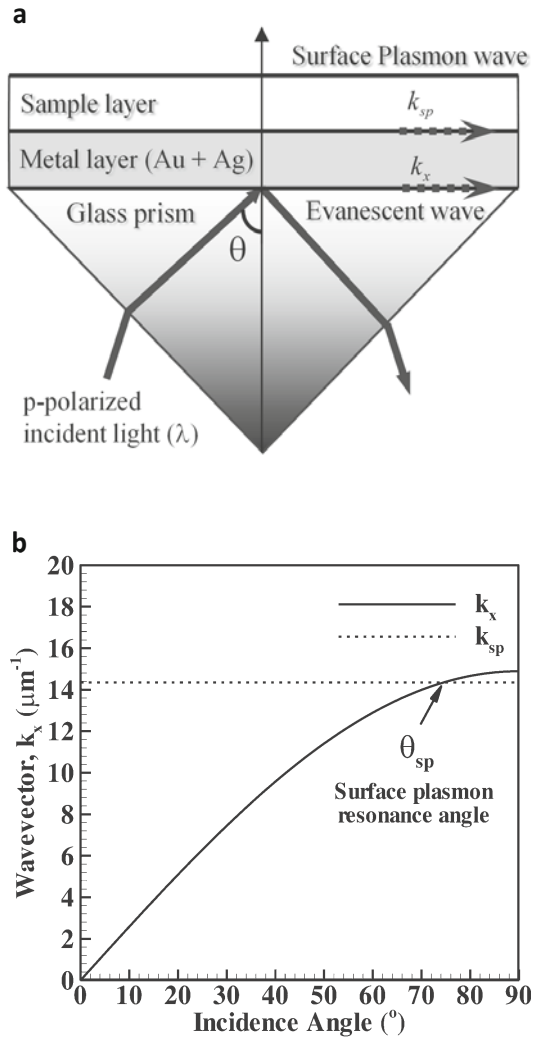


Fig. 1 Schematic of the Kretschmann SPR configuration with a metal film evaporated onto the glass prism and resonance condition. **a** Kretschmann SPR configuration. **b** Wavevectors (k_x, k_{sp}) and resonance criterion ($k_x = k_{sp}$) at a certain wavelength (633 nm)

of prism coupler based attenuated total reflection by Kretschmann (1968) and angular modulation methods are used in this study for excitation of surface plasmon waves. When surface plasmon resonance occurs the intensity of reflected light decreases and a well defined minimum of the reflectance angle is seen. The estimated results are provided for the reflectance, the phase shift, and the enhancement of magnetic field intensity in order to investigate the optical characteristics of SPR sensor and the sensitivity in finding the SPR angle.

2. Mathematical representation

There are three methods for the excitation of SPs: prism coupling, grating coupling, and waveguide coupling. Among them, the present simulation uses the Kretschmann prism coupling configuration widely adopted in most practical applications for SP excitation as illustrated in Fig. 1a. This configuration uses an evaporated metal film on a glass block with the light illuminating through the glass and causing an evanescent wave to penetrate into the metal film. Consequently, the plasmons are excited at the outer side of the film. The surface plasmon resonance is excited by totally reflected transverse magnetic (TM) light at a metal-coated interface when the wave vector of the incident light in the plane of the surface (k_x) matches with the wave vector of the surface plasmons in the metallic films (k_{sp}). The wave vector of the incident light and that of the surface plasmons in the metallic films are expressed as:

$$k_x = \frac{2\pi}{\lambda} n_p \sin\ \theta \tag{1}$$

$$k_{sp} = \frac{2\pi}{\lambda} \sqrt{\frac{\varepsilon_m \varepsilon_d}{\varepsilon_m + \varepsilon_d}} \tag{2}$$

where, ε_d and ε_m represent the complex dielectric constants of dielectrics and metal, respectively, and n_p indicates the refractive index of the prism. In addition, λ denotes the wavelength of incidence light. The matching

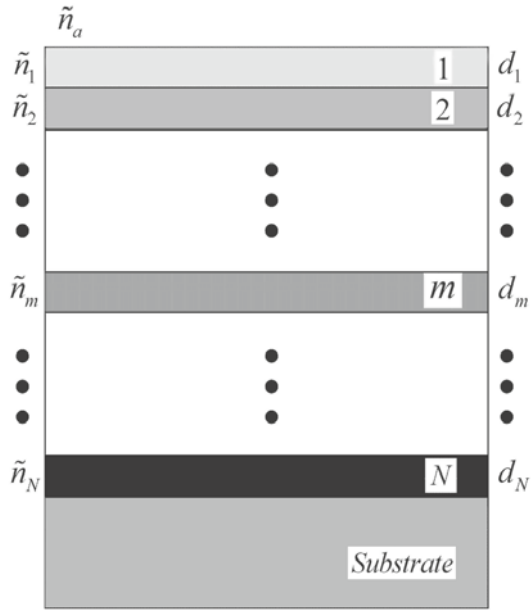


Fig. 2 Schematic of the multi-layer structure for multiple beam interference matrix (MBIM) method (M: mid-layer, N: the N-th layer, *d*: thickness of layer, and $\tilde{n}$: complex refractive index of layer)

relationship between the wave vectors k_{sp} and k_x is used to obtain the resonant angle θ_{sp}. As following Fig. 1b, the resonant angle is the angle when wave vectors k_{sp} and k_x are matched.

$$\theta_{sp} = \sin^{-1}\sqrt{\frac{\varepsilon_m \varepsilon_d}{\varepsilon_p \quad \varepsilon_m + \varepsilon_d}} \qquad \text{when} \quad k_{\text{sp}}=k_{\text{x}} \quad (3)$$

When the SPR takes place, the optical characteristics such as the reflectance, the phase shift, and the enhancement of magnetic field intensity are drastically changed. When the incident light passes through the prism–metal–dielectrics multilayer SPR structure, the light partly reflects and partly transmits in each layer. For a thick metal layer, normal reflection of the layer can be determined from Fresnel's formula, whereas in the case of a metal layer thickness, that is less than 10 times the wavelength of the incident light, the interference effect cannot be neglected. Hence, one must bear in mind the existence of wave interference in a metal layer with a mean thickness of less than 100 nm. Instead of Fresnel's formula, an alternative approach should be used.

To consider wave interference as well as variation of the complex refractive index in space, the present study uses the multiple beam interference matrix (MBIM) method derived from Maxwell's equation that relates the electric and magnetic fields. The MBIM method basically uses the characteristic matrix of a stratified medium with appropriate boundary conditions and divides the multilayer film structure into several grids as illustrated in Fig. 2. It can also estimate the reflectance, the phase shift, and the enhancement of magnetic field intensity of incident light for the lump structure. In Fig. 2, the characteristic 2 × 2 matrix M_m of *m*-th layer with a complex refractive index $\tilde{n}_m = n_m - ik_m$ and the thickness d_m is given for a transverse magnetic (TM) wave as follows:

$$\mathbf{M}_m = \begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{21} & \mathbf{M}_{22} \end{bmatrix} = \begin{bmatrix} \cos\beta_m & -\frac{i}{q_m}\sin\beta_m \\ -iq_m \sin\beta_m & \cos\beta_m \end{bmatrix} \quad (4)$$

where, β_m and q_m can be determined by

$$\beta_m = \frac{2\pi}{\lambda}\tilde{n}_m d_m \cos\theta_m \quad (5)$$

$$p_m = \tilde{n}_m \cos\theta_m \quad (6)$$

The characteristic matrix of the whole medium, considered as a pile of thin layers, is equal to

$$\mathbf{M} = \prod_{m=1}^{N} \mathbf{M}_m \quad (7)$$

The reflection and transmission Fresnel coefficients, *r* and *t*, are

$$r = \frac{\left[\mathbf{M}_{1,1} + \mathbf{M}_{1,2}\tilde{n}_s\right]\tilde{n}_a - \left[\mathbf{M}_{2,1} + \mathbf{M}_{2,2}\tilde{n}_s\right]}{\left[\mathbf{M}_{1,1} + \mathbf{M}_{1,2}\tilde{n}_s\right]\tilde{n}_a + \left[\mathbf{M}_{2,2} + \mathbf{M}_{2,2}\tilde{n}_s\right]} \quad (8)$$

$$t = \frac{2\tilde{n}_a}{\left[\mathbf{M}_{1,1} + \mathbf{M}_{1,2}\tilde{n}_s\right]\tilde{n}_a + \left[\mathbf{M}_{2,1} + \mathbf{M}_{2,2}\tilde{n}_s\right]} \quad (9)$$

Finally, the structure reflectance *R*, transmittance *T*, phase *Φ*, and enhancement of magnetic field intensity can be described as

$$R = |r|^2 \quad (10)$$

$$T = \frac{n_s}{n_a}|t|^2 \quad (11)$$

$$\Phi = \arg|r| \quad (12)$$

$$\left|\frac{H_y \;\; n/n-1}{H_y \;\; 1/2}\right| = |t_{12\cdots n}|^2 \quad (13)$$

In the present calculation, the wavelength used for incident light is 633 nm, and a water layer is taken for the sample layer as shown in Fig. 1. The optical properties of different media such as borosilicate crown optical glass (BK-7) prism, Ag, Au, and water are obtained from the literature (Silinb 1997; Weber 2003) as listed in Table 1. In addition, the thicknesses of the prism and water layers are taken as 25 mm and 30 mm, respectively for various metal film thicknesses.

3. Results and discussion

3.1 Surface plasmon resonance angle and reflectance

For the validation of the MBIM method, the present study conducts a preliminary simulation for the SPR configuration with an Ag layer on a firestone heavy glass (ZF6) prism and compares predicted reflectance with the measurement data (Shen 1998) under the conditions listed in Table 1. In Fig. 3 showing the reflectance with respect to the incidence angle, the SPR angle is estimated to be about 36.21°, and fairly good agreement with the experimental data can be seen. This suggests that the MBIM method may be used to appropriately describe the SPR characteristics.

As noted previously, the present study aims to examine numerically the optical

Table 1 Optical properties of different materials

Properties	BK-7 Prism	Gold	Silver	Water
Refractive index (*n*)	1.5168	0.13	0.27	1.332
Extinction coefficient (*k*)	0	3.16	4.18	0

Table 2 Material properties for validation with experimental data (Shen S 1998)

Layer	Thickness	Refractive index (*n*)	Extinction coefficient (*k*)
ZF-6 prism	25 mm	1.751	0
Silver	43 nm	0.099	4.001
Air	∞	1	0

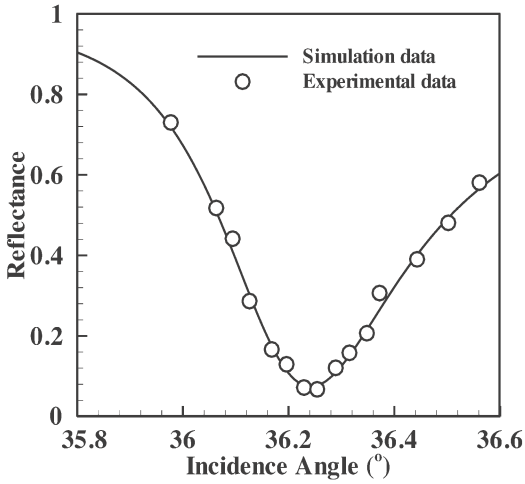


Fig. 3 Comparison of predicted reflectance with experimental data (Shen 1998)

characteristics of different SPR configurations and ultimately to find the optimum thicknesses of metal layers. This can help us understand the detailed physics behind SPR phenomena and to design more effective SPR structures with higher sensing sensitivity. In order to sensing bio-molecule interactions in a biosensor, such as antibody-antigen interactions, peptide-protein interaction, it is desirable that the SPR sensor is highly sensitive to binding interactions within the sensing region. Numerical simulations are thus conducted for two different cases: one involves a single Au film deposited on the glass prism (case 1, Au film), and the other employs a bimetallic film structure based on silver and gold films (case 2, Ag-Au film). The incidence angle varies from 0° to 90° at a fixed wavelength of 633 nm. For case 1 representing the single Au film, the predicted reflectance values are plotted for various film thicknesses as shown in Fig. 4a. From the results, an optimum Au film thickness can be found at 52 nm approximately, at which the minimum reflectance due to the presence of SPR can be found in the vicinity of 74.4°, indicating the SPR angle at 633 nm as shown in Fig. 4b.

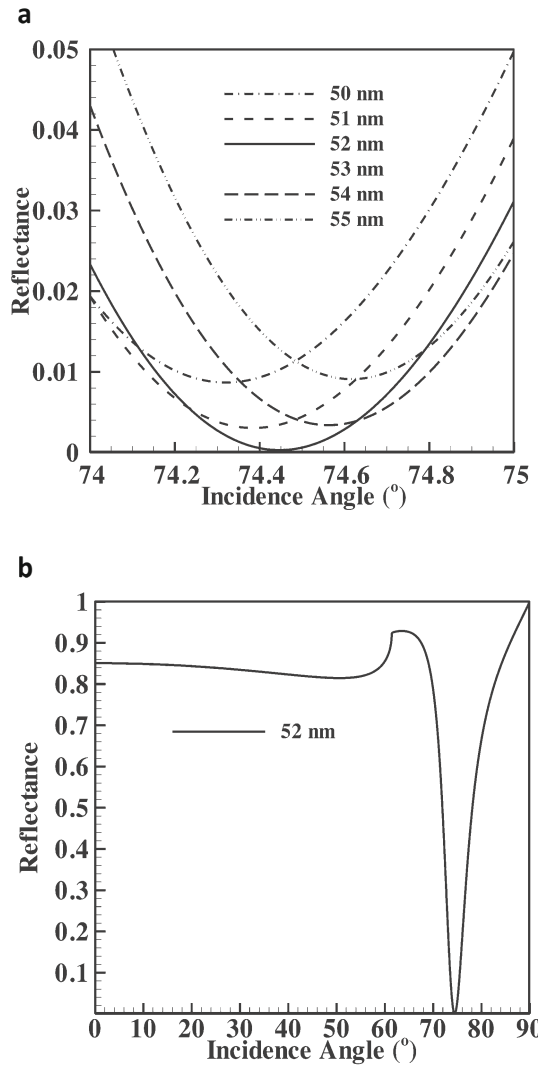


Fig. 4 **a** The predicted reflectance for various Au film thicknesses ranging from 50 nm to 55 nm (case 1). (magnification from 74° to 75° incidence angle). **b** Reflectance for 52 nm Au film thickness

Meanwhile, for the bimetallic SPR structure (case 2), it can be seen that the sensitivity of reflectance in the SPR structure decreases with increasing Au film thickness. In fact, a minimum gold film thickness is limited to 5 nm in the present calculation when considering the fabrication limit using vacuum thermal evaporating or the metal sputtering process. From Fig. 5a it is found that an op-

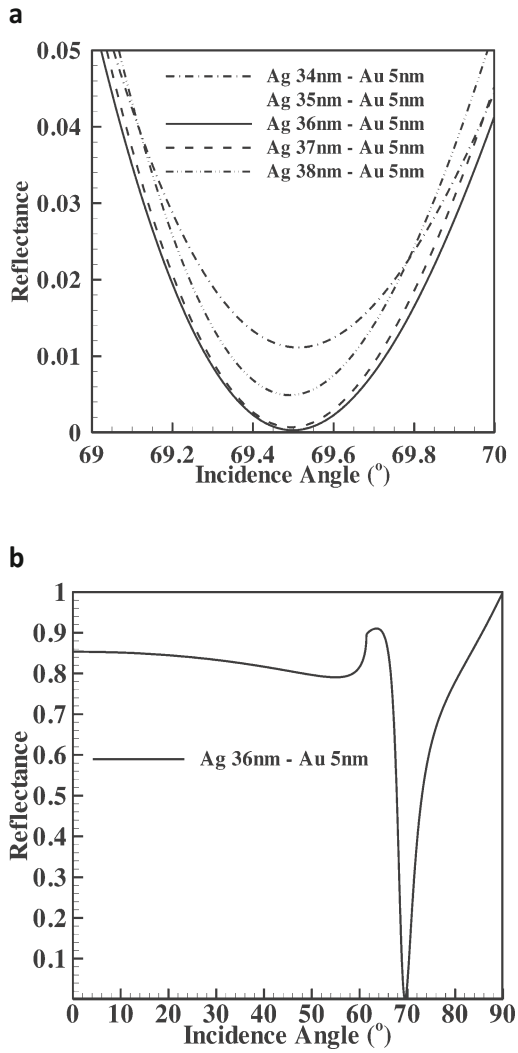


Fig. 5 Reflectance for various Ag film thicknesses at fixed 5 nm Au film thickness in the bimetallic SPR structure (34 nm ~ 38 nm). **a** Reflectance for various Ag film thicknesses ranging from 34 nm to 38 nm (magnification from 69° to 70° incidence angle). **b** Reflectance for 36 nm Ag film thickness

timum thickness for a silver film is estimated to be 36 nm for case 2 and the corresponding SPR angle for the bimetallic structure is about 69.49°.

3.2 Comparison of sensitivity for single and bimetallic structures

To acquire such unique SPR signals occurring at the interface with higher sensitivity in SPR biosensor applications, it is necessary to compare the sensitivities of several sensing parameters such as the reflectance, the phase shift and the enhancement of magnetic field intensity which are closely associated with the SPR phenomena. Hence, we compare the sensitivity of reflectance, phase shift, and enhancement of magnetic field intensity by evaluating the full width at half maximum (FWHM). The FWHM is the parameter commonly used to describe the width and sharpness of a curve of function or data and is given by the difference between two extreme values of the independent variable when the dependent variable is equal to half of its maximum value. The numerical results are compared for different structures (case 1 and case 2) in Fig. 6–8, indicating that for case 1 and case 2 the FWHM of reflectances is estimated to be 6.21° and 5.24° respectively. The values as a function of other parameters are summarized in Table 3.

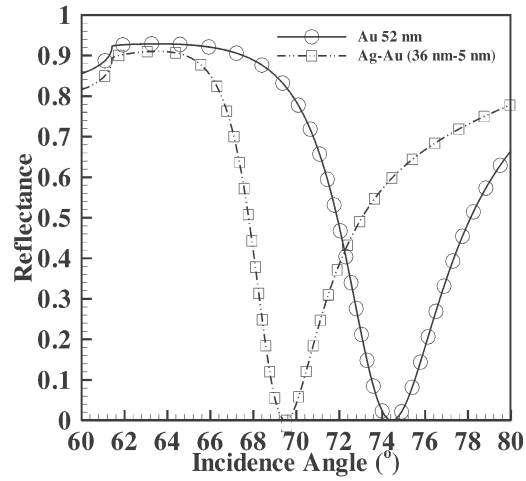


Fig. 6 Reflectance of optimum single Au and bimetallic Ag-Au SPR sensor

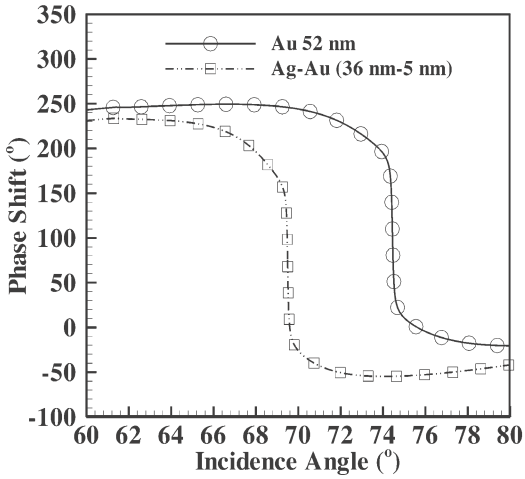


Fig. 7 Phase shift of optimum single Au and bimetallic Ag-Au film SPR sensor

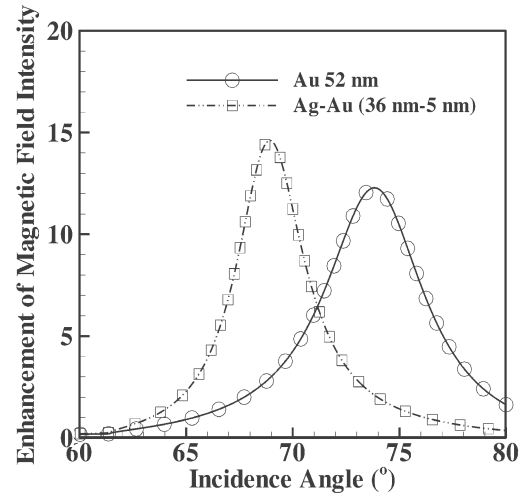


Fig. 8 Enhancement of magnetic field intensity of optimum single Au and bimetallic Ag-Au film SPR sensor

From the results, it can be seen that the estimated FWHM of reflectance becomes larger than that of the magnetic field intensity as well as the phase shift. A larger FWHM makes it difficult to distinguish small changes in the peak value. The deeper and narrower the resonance signal is, the more accurate the determination of SPR angle. Considering that a smaller value of FWHM indicates higher sensitivity of SPR, it can be concluded that the measurement of phase shift or enhancement of magnetic field intensity is recommended in acquiring more sensitive SPR signals compared to reflectance, widely adopted in earlier published works. In particular, a bimetallic SPR structure would be a more suitable way to obtain higher sensitivity of SPR sensors compared to a single film structure (case 1).

Conclusions

The present study investigates the optical characteristics of single and bimetallic SPR structures by using the MBIM method, and provides numerical predictions of reflectance, phase shift, and enhancement of magnetic field intensity. The MBIM method used in the present study can predict the optical characteristics of multilayer structures effectively, suggesting that it is appropriate for describing SPR characteristics. In addition, the optimum thicknesses of single Au and bimetallic Ag-Au film structures are estimated and their values are useful for designing SPR structures with high sensitivity. From the comparison of the sensitivity for two different SPR structures, the measurement of phase shift or enhancement of magnetic field intensity is recommended in acquiring more sensitive SPR

Table 3 FWHM value of reflectance, phase shift, and enhancement of magnetic field intensity of single Au and bimetallic Ag-Au film structures

Configuration	Value of full width half maximum (°)		
	Reflectance	Enhancement of magnetic field intensity	Phase shift
Single Au film (case 1)	6.21	5.51	0.28
Bimetallic Ag-Au film (case 2)	5.24	3.82	0.39

signals because of the relatively small value of the reflectance's FWHMs. Comparing between single and bimetallic SPR structures, it can be concluded that bimetallic SPR structures using different metal layers represent a more suitable way to obtain higher sensitivity of SPR sensors compared to a single film structure (case 1). In that case, the estimated sensitivity of the phase shift is ten times better than other physical parameters, such as reflectance and enhancement of magnetic field intensity.

Abbreviations

FWHM	full width half maximum
MBIM	multiple beam interference matrix
SPs	surface plasmons
SPR	surface plasmon resonance
SPW	surface plasmon wave
TM	transverse magnetic

References

Chen KH, Yang SL, Chang KC (2007) Measurement of small differences in refractive indices of solutions with interferometric optical method. Opt Laser Eng 45: 1071–107

Cui Z, Wang X, Li Y, Tian GY (2007) High sensitive magnetically actuated micro-mirrors for magnetic field measurement. Sensor Actuat A-Phys 138: 145–150

Göktepe M (2001) Non-destructive crack detection by capturing local flux leakage field. Sensor Actuat A-Phys 91: 70–72

Ho HP, Yuan W, Wong CL, Wu SY, Suen YK, Kong SK, Lin C (2007) Sensitivity enhancement based on application of multi-pass interferometry in phase-sensitive surface plasmon resonance biosensor. Opt Commun 275: 491–496

Homola J, Yee SS, Gauglitz G (1999) Surface plasmon resonance sensors: review, Sensor Actuat B-Chem 54: 3–15

Joubert PY, Pinassaud J (2006) Linear magneto-optic imager for non-destructive evaluation. Sensor Actuat A-Phys 129: 126–130

Kretschmann E, Raether H (1968) Radiative decay of non-radiative surface plasmons excited by light. Z Naturforsch A 23: 2135–2136

Lee JY, Shih HC, Hong CT, Chou TK (2007) Measurement of refractive index change by surface plasmon resonance and phase quadrature interferometry. Optics Communications 276: 283–287

Nylander C, Liedberg B, Lind T (1982) Gas detection by means of surface plasmon resonance. Sensor Actuat 3: 79–88

Liedberg B, Nylander C, Lundstrøm I (1983) Surface plasmon resonance for gas detection and biosensing. Sensor Actuat 4: 199–304

Liedberg B, Nylander C, Lundstrøm I (1995) Biosensing with surface plasmon resonance – how it all started. Biosens Bioelectron 10: i–ix

Shen S, Liu T, Guo J (1998) Optical phase-shift detection of surface plasmon resonance. Appl Optics 37: 1747–1751

Silinb V, Weetallc H, David JV (1997) SPR studies of the nonspecific adsorption kinetics of human IgG and BSA on gold surfaces modified by self-assembled mono-layers (SAMs). J Colloid Interf Sci 185: 94–103

Tian GY, Sophian A, Taylor D, Rudlin J (2005) Multiple sensors on pulsed eddy-current detection for 3-D subsurface crack assessment. IEEE Sens J 5: 90–96

Weber MJ (2003) Handbook of optical materials. CRC Press, pp 55–69

Deflection-based flow field sensors – examples and requirements

27

Sebastian Große, Wolfgang Schröder

Contents

Abstract

The assessment of near-wall fluid motion, acoustic vibrations, or wall-shear stress is essential in many biological and engineering systems. This necessity is evidenced by a huge diversity of sensing devices in both fields. While nature shows a broad diversity of filiform hair-like fluid sensing devices[1] which have been improved towards perfection over millions of years, researchers have recently attempted to copy and adapt these biological examples for technical applications. In particular the fish lateral line flow sensor and the filiform arthropod hair-like medium motion sensors have inspired researchers to develop artificial hair sensor arrays based on flexible cantilevers and micro-posts. Therefore, in this article, after a brief description of some examples of biological deflection-based sensor devices, recent man-made cantilever-based sensors for the detection of near-wall fluid motion and wall-shear stress will be discussed. Furthermore, the advantages and disadvantages of different designs and general fluid-mechanical and technical requirements for flow sensors will be outlined.

1. Biological examples for deflection-based flow field sensing

Nature provides a whole diversity of flow-sensing devices many of which are based on cantilever-type hair receptors. The lever arms of these sensors tilt or bend in re-

Sebastian Große
RWTH Aachen, Institute of Aerodynamics and Chair of Fluid Mechanics
Wüllnerstr. 5a, 52062 Aachen, Germany
e-mail: sebastian.grosse@rwth-aachen.de

[1] Please note that "filiform hair-cell sensing device" is a misnomer ignoring the difference between a filiform hair or a hair-like sensor and a hair cell which is a well defined biological structure.

sponse to fluid forces, causing mechanical stresses in sensing neurons to which they are connected. Note, throughout the text, we will use the word "bending", if the sensor structure is deflected by internal elastic deformation. If a non-flexible structure reacts to the fluid forces with a movement without internal bending, we will call this movement "tilting". The medium flow sensing hairs help animals to feel obstacles in their vicinity, detect surrounding fluid motion, warn the animal of potential predators or help in sensing and locating prey.

An example of such sensors are the trichobothria on the walking leg of the spider *Cupiennius salei*, which serve as extremely sensitive detectors of the slightest air motions in the spider's vicinity (Humphrey et al. 1993, Humphrey and Barth 2008). The complex nature of these sensor devices allows the spider to precisely detect the motion of prey and to successfully differentiate between the signature of potential food from random wind motion, thereby saving the spider's energy resources. Barth et al. (1993) studied the dynamics of these hairs and showed that the spider took advantage of resonances of the hair devices at a certain frequency range, making each individual sensory hair geometry especially sensitive to air fluctuations at these frequencies. In other words, at certain frequencies, the tilting or bending of the device is stronger compared to all other frequencies. It is important to note, however, that in the overwhelming majority of cases the ratio between air and hair-tip movement is smaller than one. For their investigations, the authors placed hairs with lengths varying between 400 and 1 150 µm in an oscillating air medium. A detailed discussion of the results from this analysis of the dynamic properties of the hairs is given in Barth et al. (1993). The findings indicated that the sensory hair is not only especially sensitive to air fluctuations at a certain frequency range but, furthermore, Humphrey and Barth (2008) showed that the mechanical system reacted most sensitively to changes in the velocity fluctuations in this frequency range, filtering other frequencies (Barth and Höller 1999) to the extent that the receptor reacts like a band-pass filter at the preferred frequency (Humphrey and Barth 2008). Combinations of hairs with different geometrical and mechanical properties allow coverage of a large frequency range, so that the interesting frequency spectrum at which potential prey in the spider's immediate vicinity that cause air fluctuations is fully covered by the receptors of the spider. The high sensitivity of the trichobothria allows the spider to detect even the slightest fluid motions in its vicinity. A quantitative near-wall fluid flow or wall-shear stress measurement would inevitably be spoiled once a narrowly tuned sensor experiences an excitation at its resonance frequency. Comprehensive reviews concerning the mechanics of these biological sensor structures are given by Humphrey and Barth (2008) and Shimozawa et al. (1998).

Rat's whiskers should serve as another example for a biological cantilever-based sensor (see also Chapter IV,2 by RB Towal et al.). Rats possess a bunch of whiskers located around their nose and the animal moves them rhythmically back and forth to tactually capture information about their environment (Krupa et al. 2001). The whiskers are arranged in a regular array on the rat's face, with each whisker inserted into a densely innervated follicle filled with many different types of mechano-receptors. This makes the whiskers extremely powerful as a tool for the rat to obtain considerable information about an object's spatial properties, including its size, shape, orientation, and texture.

Non-motile (or primary) cilia functioning as fluid shear-stress sensors in endothelial cells of the chicken embryonic endocardium have been identified by van der Heiden et al. (2006). These nearly cylindrical cilia are only

a few µm long. They start to disassemble when subjected to high shear-stress levels. In areas in which moderate and low shear-stress levels are present, these sensors are believed to detect fluid shear-stress pattern changes, which go hand in hand with the cardiovascular development. The authors suggest that the shear stress plays an essential role in the chicken's cardiovascular development by forcing the gene expression in endothelial cells. The exact mechanism by which endothelial cells sense shear stress is subject to further investigation.

The lateral line sensor will serve as a final illustration of a natural cantilever-based mechano-receptor (see also Chapter VI,6 by N Izadi and GJM Krijnen). Fish use so-called hair cells to detect weak water motions and local pressure gradients. The lateral line system consists of neuromasts (which contain the hair cells) placed on the body surface or in lateral line canals. Fish use the information from the lateral line system not only for the detection and identification of prey and predators but also as sensor input for its own spatial orientation and navigation. The working principle and the physiology of the lateral line system of fish has been investigated in depth, for example by Dijkgraaf (1962), Engelmann et al. (2000) or Bleckmann (2007). More detailed information on the lateral line system of fish can be found in Chapter I,2 by H. Bleckmann.

Although the above list is far from complete it illustrates the broad diversity of biological cantilever-based sensors which find applications in the detection and measurement of fluid motion or of acoustic vibrations. The examples have shown that in many cases medium flow-sensing hairs assist the animal's predation and escape behavior or simply serve as sensors for their orientation in the surrounding fluid. Similar sensory systems for the reception of medium flow are found in many animal groups. Although the different devices may seem quite similar, medium-flow sensors differ in details that reflect their adaptation to the various boundary conditions under which they are working. Nevertheless, the broad diversity of the cantilever-based sensing principle in the animal kingdom is self-explanatory and indicative of its success.

2. Artificial devices for deflection-based flow field sensing

From the discussion above it is not surprising that many attempts have been made to copy nature's genius and to develop and construct artificial cantilever-based sensors. As in the natural examples, the fields of application are diverse, ranging from artificial lateral line organs for the navigation of, for example, unmanned vehicles, the assessment of flow rate, or applications such as the measurement of the local wall-shear stress, an essential parameter in many fluid-mechanical applications and in turbulence research. In the following, some recent designs are briefly described.

Ozaki et al. (2000) developed different hair-like sensor designs. An initial design consisted of flat, rectangular micro-cantilevers of different lengths ranging from 400 µm to 1 200 µm, a width of 230 µm and a depth of 10 µm with strain gauges at the bottom. This geometry results in very different moments of inertia along the different directions, such that only uni-directional fluid forces could be resolved. The authors related the deflection of the cantilevers measured by the strain gauges to the local velocity and calibrated the sensor statically in a wind tunnel flow. A second sensor design consists of a 3 000 µm long cylindrical structure suspended by four cantilever springs. Strain gauges mounted on

the cantilever springs detect the individual deflection and tortion of these cantilevers as a measure of the fluid forces exerted on the cylindrical structure and, hence, of the local velocity. Due to the symmetric shape of this setup, the sensor possesses a multi-directional sensitivity. An issue not addressed was whether the sensor design suffers from cross-axis sensitivity due to a superposition of primary-axis deflection and cross-axis torsion of the cantilever-springs.

Another design for a flow cantilever has been proposed by Fan et al. (2002) and Chen et al. (2003). A vertical cantilever is mounted on a horizontal lever, forming an L-shaped structure. The size of the vertical cantilever beam is approximately 1 mm and protrudes into the local flow field. The horizontal lever and the vertical structure are connected via a rigid link to the extent that the tilting motion of the vertical beam is transferred to a piezo-resistive strain gauge placed at the base of the horizontal lever. The structure is manufactured using a so-called plastic deformation magnetic assembly (PDMA) method, a technique described in more detail in Zou et al. (2001). Magnetic forces are used to bend the surface micro-machined beams into the desired vertical position. Due to the plastic deformation, the beams remain in the bent position. Since single cantilevers can only detect uni-directional velocities, the authors grouped arrays of sensors with different frontal orientations.

Tucker et al. (2006) describe a sensor structure based on a cylindrical micro-post with a diameter of 80 µm and a length of 600 µm rigidly mounted on a horizontal lever bar. SU-8 epoxy is used as cylinder material. The high rigidity of the SU-8 material causes the entire cylinder to be tilted by the fluid forces This tilting deflection causes the flexible horizontal lever to bend, which can be measured by silicon piezo-resistive strain gauges at its base. With the geometry of this read-out device, the sensor allows the determination of only one flow direction. To detect the two-dimensional flow field distribution, the authors use neighboring pairs of sensors with strain gauges oriented orthogonally to each other in the same plane. The sensor achieves a flow rate sensitivity of 0.6 mm/s before reaching the electronic and intrinsic noise level. The resonance frequencies of beams at the specific mechanical and geometric dimensions are considered to be in the order of a few kHz. Further studies with this sensor applied as an artificial lateral line organ have been reported by Yang et al. (2010). The authors used a generic hydrodynamic (dipole source) and, furthermore, a realistic stimulus (flicking crayfish-tail) to investigate the capability of the sensor design in correctly localizing the source of stimulus.

Engel et al. (2006) and Chen et al. (2007) presented an artificial hair-like sensor based on cylindrical polyurethane structures, and further details about this sensor can be found in Liu (2007). The authors positioned the posts on commercial conductive, circularly aligned, polyurethane force sensitive resistors (FSR) to detect the deflection of the post structure and to allow for a multi-directional sensitivity. Due to the conductive read-out of the cantilever tilting, even the cylindrical version of their posts showed a remaining on to off-axis sensitivity ratio of 14.2 dB. The authors, aiming for uni-directional sensitivity, improved the sensor geometries by increasing the lateral stiffness, so that the modified sensors showed the desirable off-axis mechanical insensitivity. However, this also impeded the possibility of multi-directional flow detection.

A hair-like sensor to measure acoustic pressure disturbances and flow fields is reported by Dijkstra et al. (2005) and Krijnen et al. (2006). The authors use SU-8 cricket-sensory hairs to detect drag forces on the sensor structure. With sensor heights of up to 1 mm the authors protrude deep into the

near-wall flow field achieving a high sensitivity in the device. The tilting deflection is detected capacitively at the sensor-hair base using different kinds of suspensions. The authors used arrays of similar structures to increase the total capacitive response, i. e., increasing the sensibility of the structures. The authors have studied very extensively the dynamic response of the structure and its directional sensitivity.

Finally, a micro-pillar shear-stress sensor has been discussed extensively in the literature (e. g. Grosse 2008, Grosse and Schröder 2009 a). This sensor design consists of cylindrical micro structures made of the highly flexible silicone polydimethylsiloxane (PDMS). Different to the designs described before, the vertical cantilever structure itself is deflected. The sensors are up to approximately 1 mm in height and possess aspect ratios (length vs. diameter) of up to L/D=20. Images of the sensor are given in Fig. 1. The pillar-tip deflection caused by the exerting fluid forces is measured optically. Due to the symmetric cylindrical shape, the micro-pillar sensors possess a constant stiffness along all radial directions.

The sensor design is described in detail in Grosse and Schröder (2009 a) and Grosse (2008). The discussion in the first of these papers addresses material characteristics, sensor-structure related error sources, various sensitivity and distinct sensor performance questions as well as presenting a detailed discussion of pressure-sensitivity aspects. In Grosse et al. (2008) an experimental dynamic calibration of the sensor structure is described. The results indicate that typical sensor structures with the dimensions described above possess low pass filter characteristics in water media with damped eigenfrequencies in the order of several hundred Hz. In air, the sensor structures show a strongly resonant behavior – due to the reduced added mass and friction in this medium – with resonance frequencies in the order of a few thousand Hz. Application of micro-pillars to detect turbulent wall-shear stress in water flows are discussed in Grosse and Schröder (2008 a, 2008 b, 2009 b, 2009 c). Grosse and Schröder (2009 d) also discuss the feasibility of wall-shear stress measurements in turbulent wind-tunnel flow with the sensor technique.

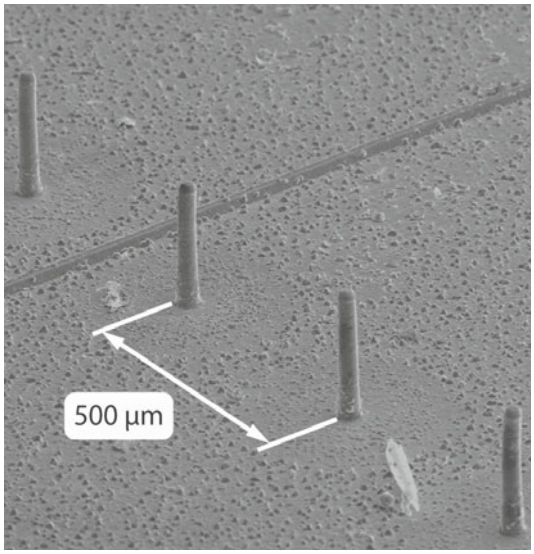


Fig. 1 Scanning-Electron Microscope (SEM) images of a pillar line

3. MEMS vs. MOMS

The discussion above shows the great diversity of data acquisition techniques. These so-called Micro-Electro-Mechanical Systems (MEMS) transform the mechanical reaction of the sensor to the exerted forces into an electrical signal, i. e., a voltage, by capacitive, conductive or (piezo)-resistive means.

This principle has a couple of advantages. A prominent feature is the possibility of designing the sensor structures reasonably

rigidly with, accordingly, high resonance frequencies since even slightest deflections can be detected with sufficient sensitivity by the mechano-electric coupling. Similarly, sensor structures can be kept very small to the extent that they can be applied successfully even at high Reynolds numbers.

However, the MEMS-technique also has some disadvantages. The read-out principle requires the implementation of secondary structures on the sensor base. One can think of the necessary cantilever springs as mechanical read-out devices such as strain gauges or simply electrical supply wiring. Many of the 1-DOF and 2-DOF devices (degree of freedom) suffer from cross-axis sensitivity, that is, they are sensitive to forces exerted along the axis perpendicular to the primary sensory axis. For multi-directional devices, which detect the two-directional deflection of cantilever beams with base-mounted mechanical receptors, e. g., strain gauges or piezo-resistive devices, the arrangement of these elements might become sensitive to deflections along perpendicular directions, making an identification of the originating force direction difficult or even impossible. Cross-axis sensitivity can be minimized by increasing the stiffness of the sensor structure along the perpendicular direction so that the parasitic off-axis contributions are negligible. It goes without saying that this also implies the non-capability of actually detecting fluid motion perpendicular to the primary sensor axis, which is an essential feature in turbulent, non uni-directional flows. To overcome this issue MEMS devices have been assembled with different sensitivity orientation. However, this causes the flow field not to be measured in a single position and the combination of sensor signals to reconstruct the medium flow in direction and strength needs to be performed with great care (Suzuki 1992).

The micro-pillar sensor described, for example, in Grosse and Schröder (2009a) represents a Micro-Opto-Mechanical Systems (MOMS) device, transferring the mechanical deflection into an optical signal. Due to the optical detection principle there is no need for further structure on the wall, and as such, the assessment of the two-dimensional wall-shear stress distribution at high spatial resolution is only limited by the minimum distance between the pillars required to ensure neighboring pillars not to interfere with each other from a fluid mechanical or electrical point of view. The pillar has a symmetrical structure with no preferred sensing direction, that is, the deflection is a direct representation of the direction and amplitude of the exerted force. Additionally, the optical detection principle does not suffer from cross-axis sensitivity, making measurements of the near-wall flow field vectors possible.

However, these advantages are also accompanied by a couple of drawbacks. The optical detection requires optical access which many applications can not be granted and which restricts the use of the sensor to mainly scientific applications. Due to the limited maximum optical resolution of standard camera-lens systems, the pillar structures need to be flexible enough to allow a proper evaluation of the pillar deflection. Due to the limited size of optical sensors (CCD, CMOS), field-of-view and optical resolution are mutual characteristics requiring a compromise solution. In the case of the micro-pillar sensor, the optical principle is very similar to that of Particle-Image Velocimetry (PIV), to the extent that similar assumptions apply for the setup of the micro-pillar sensor. The PIV-technique is further described in Raffel et al. (2007).

The necessarily high flexibility of the structures makes the resonance frequency of the sensor comparably lower than those of many other common MEMS-cantilever devices. This reduces the achievable dynamic range and restricts the application of the

sensor to low and moderate Reynolds numbers.

Another drawback of the optical recording principle is the limitation of the available amount of memory in conventional digital cameras. Generally, this aspect reduces the number of recordable images to a few thousand samples, complicating proper statistical evaluation of data especially at high recording frequency.

While MEMS devices emit an electrical voltage, which can be directly processed and evaluated, the optical detection principle produces raw images, which require a very time-consuming image evaluation (again, similar to that of Particle-Image Velocimetry, PIV) before the actual fluid-flow data is obtained. Hence, the use of the optical evaluation technique impairs the use of the micro-pillar sensor in flow control application, which would require a real-time evaluation of the data.

However, the scientific field of application, for which the micro-pillar sensor was originally intended, easily copes with most of these stated drawbacks. Finally it should be mentioned that besides the general capability of PDMS as sensor material – a detailed discussion of which is given in Grosse (2008) and Grosse and Schröder (2009a) – it has proven to be very robust to the daily user and has endured some rough handling; in contrast most common MEMS-sensors are very brittle making daily use problematic.

4. Fluid mechanics of the sensing principle and sensor requirements

We have so far discussed different technical realizations of cantilever-based sensors for the detection of near-wall fluid motion. In the last paragraph, recent designs have been discussed based on the principle of detection, i. e., the present designs have been divided into two categories, one where the sensor deflection is measured electrically and another where the sensor tilting or bending is detected optically. On a general level, the pros and cons of both concepts have been discussed, based on the designs available in the literature. A detailed discussion of the data acquisition and processing is, however, beyond the scope of the present article, and interested readers are referred to the very comprehensive discussions in the work available in the literature – the relevant papers have been presented in section 3. For the micro-pillar sensor, detailed information on the image acquisition and image processing can be found in Grosse (2008).

In this section, we will focus on the fluid mechanics of the sensory hairs for a better understanding of the sensor's working principles and to be able to judge the general feasibility of a qualitative and quantitative detection of the fluid field. At this point, we will concentrate solely on the mechanics of the fluid-sensor coupling, with the discussion focusing on turbulent flows, since most sensors are applied in such a flow field.

All cantilever-type receptors share a filament structure that protrudes into the near-wall flow field. Due to the fluid forces acting on the structure it is bent or tilted and the deflection of the sensory hair is measured. It is important to note, as mentioned in the last section, that for some designs it is not the hair shaft itself that bends but rather the lever devices on which it is mounted (caused by the hair shaft, which is tilted by the fluid flow). However, in terms of fluid-mechanics the concepts can be considered similar.

Most of the structures discussed in this paper are characterized by their small dimensions which range in the order of a few µm to mm, and most fluid flow characteris-

tics from the data in the literature are such that the local flow field around the structure is at very low Reynolds numbers (Re ≤ 1) and can be considered creeping. Here, the Reynolds number is based on the local velocity *U*, the fluid kinematic viscosity *v*, and a characteristic cross-sectional dimension of the sensor (e. g. its diameter).

Under turbulent conditions, the total load on the structure is not only caused by the drag forces acting on the structure due to the fluid flowing around it, but also pressure forces due to local pressure fluctuations which need to be considered. In pressure-driven flows a constant pressure difference acts on the structure along the streamwise direction. Grosse and Schröder (2009 a) have demonstrated that in most cases these pressure contributions can be neglected compared to the drag forces. Note, however, that even for Oseen flow around the sensor structures at low Reynolds number, the drag force on the structure has shear and pressure contributions. These, however, are already accounted for by experimental calibrations or analytical drag coefficients. At larger Reynolds numbers, i. e., Re ≥ 1, local flow separations of the flow around the structures might cause additional pressure forces to more severely contribute to the sensor deflection. However, we will further assume that the local Reynolds number is small enough, i. e., Re ≤ 1, to the extent that the flow field nearly perfectly follows the structural contour.

The drag forces, as a function of the local fluid velocity, act as a distributed load on the sensor structure causing the hair shaft deflection. To be more precise, the sensor bending or tilting is a function of the integral contribution of the fluid field along the structure, and, consequently, the velocity distribution along the sensor has an important impact on the magnitude of the sensor deflection. The knowledge of the flow field is hence a necessary prerequisite for inferring any quantitative information from the sensor deflection. For example, imagine an inflection point of the velocity profile along the sensor lengths, so that the fluid forces would be compensated on their integral mean. In this case it is necessary to have a detailed knowledge of the flow field around the sensor structure.

While at the wall, the fluid velocity needs to become zero, following Newton's no-slip condition. The velocity profile at some distance from the wall can range from nearly linear as in the case of the viscous sublayer of turbulent flows, parabolic as in the case of steady, fully developed laminar flow to bulgy velocity profiles following complex Bessel's functions as in the case of oscillating flow fields. Following the discussion above, each of these flow fields will contribute differently to the sensor drag and, hence, its deflection. Alfredsson et al. (1988) and Grosse and Schröder (2009 a) have discussed, how under certain conditions the velocity distribution in the vicinity of the wall can be considered linear – on average and instantaneously – so that a proportional relationship between the sensor deflection and the integral velocity profile is defined. Note in the case of the micro-pillar sensor, the linearity is necessary since the sensor structure is statically calibrated in the linear velocity profile of a plate-cone rheometer flow (Grosse et al. 2006). The similitude of the flow condition during calibration and actual measurement allows for an easy quantative transfer of the calibration results to the actual measurement. It is important to note, that the region, in which the near-wall velocity profile can be considered linear, is limited to a few hundred µm in most turbulent flows, thereby imposing a strong restriction to the maximum allowable sensor length. Experimental and numerical results show that the near-wall turbulent velocity profile can be assumed linear up to $y^+ = 5 - 6$. Herein, $y^+ = \nu \cdot y/u_\tau$ is the non-dimensional wall-distance in vis-

cous units with ν being the kinematic viscosity of the fluid and u_τ the friction velocity. The wall region at which the instantaneous velocity profile can be considered linear is even narrower. A detailed discussion can be found in Grosse and Schröder (2009a). In oscillating flow, laminar or turbulent, the situation becomes even more complicated and the reader is referred to the work of Womersley et al, who thoroughly addressed this issue (e. g. Womersley 1955).

Apart from the static calibration, sensor dynamics need to be critically discussed. It is not enough to only look at the sensor dynamics alone, but rather the interaction of the surrounding fluid needs to be taken into account. For biological sensors a detailed discussion can be found in the work of Humphrey and Barth (2008) and Shimozawa et al. (1998), but it can be easily transferred to technical concepts. A discussion of this issue can also be found in Grosse et al. (2008), where the authors also describe a dynamic calibration for the micro-pillar shear-stress sensor. The authors show that added mass effects of the ambient fluid and fluid-structure friction strongly influence the sensor dynamics. In the case of the micro-pillars the sensor transfer function ranges from a strong resonance in the case of air as fluid to a damped transfer function in the case of water or oil.

Besides the above requirement many other features need to be addressed. Most of the requirements are controversial: A high dynamic bandwidth competes with a reasonable sensitivity. The use of large surface areas opposed to the fluid motion increases the sensor response but also leads to an increased spatial averaging of the near wall flow field, resulting in a deterioration of the detected dynamic characteristics.

Earlier it was mentioned that field-of-view and local spatial resolution are controversial parameters, especially for MOMS devices. However, similar restrictions also apply to MEMS devices where the handling of several hundred sensors mounted on small surface areas leads to electrical coupling and problems in the parallel signal processing and acquisition.

Finally, the non-intrusiveness, that is, the influence of the sensor on the development of the flow field, needs to be negligible. In the case of cantilever-based sensors, in particular, in turbulent flows the height of the viscous sublayer limits the maximum allowable size of the structure with strong limitations to the achievable static and dynamic response.

In conclusion, to properly detect the flow field in the vicinity of surfaces the complete spectrum of fluid velocity should at least be capable of being properly captured. Additionally, the sensor needs to possess sufficient dynamic bandwidth to properly detect even the highest frequency part of the flow. Consequently, the sensor dimension must be small enough to be able to actually capture the associated scales without spatial integration.

For a more detailed discussion of these aspects the reader is referred to Grosse and Schröder (2009a).

Conclusions

Nature's diversity of fluid sensing devices has inspired many scientists to develop artificial cantilever-based sensors, which allow the detection of near-wall fluid-flow fields by measuring the deflection of artificial hair-like devices by the acting fluid forces. In a discussion of sensor designs reported in the literature, advantages and disadvantages of electro-mechanical and opto-mechanical concepts have been contrasted. The discussion in the last section of this article has shown that the choice of an appropriate sensor is subject to a large variety of conditions. Constraints arising from fluid mechanical restrictions, sensor-sensitivity based requirements, and structure-mechanical prerequisites, make such a decision extremely difficult but also scientifically challenging.

References

Alfredsson PH, Johansson AV, Haritonidis JH, Eckelmann H (1988) The fluctuating wall-shear stress and the velocity field in the viscous sublayer. Phys Fluids 31(5): 1026–1033

Barth FG, Höller A (1999) Dynamics of arthropod filiform hairs V. The response of spider trichobothria to natural stimuli. Phil Trans R Soc Lond B 354: 183–192

Barth FG, Wastl U, Humphrey JAC, Devarakonda R (1993) Dynamics of arthropod filiform hairs II. Mechanical properties of spider trichobothria (*Cupiennius salei* Keys.). Phil Trans R Soc Lond B 340: 445–461

Bleckmann H (2007) The lateral line system of fish. In: Hara T and Zielinski B (eds) Sensory systems neuroscience. Elsevier Academic Press, Amsterdam, Netherlands; Boston, MA, 411–453

Chen J, Fan Z, Zou J, Engel J, Liu C (2003) Two-dimensional micromachined flow sensor array for fluid mechanics studies. J Aerosp Eng 16(2): 85–97

Chen N, Tucker C, Engel J. M, Yang Y, Pandya S, Liu C (2007) Design and characterisation of artificial haircell sensor for flow sensing with ultrahigh velocity and angular sensitivity. J MEMS 16(5): 999–1014

Dijkgraaf S (1962) The functioning and significance of the lateral-line organs. Biol Rev 38(1): 51–105

Dijkstra M, Baar JJ, Wiegerink RJ, Lammerink TSJ, Boer JH, Krijnen GJM (2005) Artificial sensory hairs based on the flow sensitive receptor hairs of crickets. J Micromech Microeng 15: 132–138

Engel J, Chen J, Liu C, Bullen D (2006) Polyurethane rubber all-polymer artificial hair cell sensor. J MEMS 15(5): 729–736

Engelmann J, Hanke W, Mogdans J, Bleckmann H (2000) Neurobiology – hydrodynamic stimuli and the fish lateral line. Nature 408(6808): 51–52

Fan Z, Chen J, Zou J, Bullen D, Liu C, Delcomyn F (2002) Design and fabrication of artificial lateral line flow sensors. J Micromech Microeng 12(5): 655–661

Große S (2008) Development of the micro-pillar shear-stress sensor MPS³ for turbulent flows. Doctoral thesis, Faculty of Mechanical Engineering, RWTH Aachen University, Shaker Verlag, Aachen

Große S, Schröder W (2008a) Dynamic wall-shear stress measurements in turbulent pipe flow using the micro-pillar sensor MPS³. Int J Heat Fluid Flow 29(3): 830–840

Große S, Schröder W (2008b) Mean wall-shear stress measurements using the micro-pillar shear-stress sensor MPS³. Meas Sci Technol 19(1): 015 403

Große S, Schröder W (2009a) The micro-pillar shear-stress sensor MPS³. Sensors 9(4): 2222–2251

Große S, Schröder W (2009b) Two-dimensional visualization of turbulent wall-shear stress using micro-pillars. AIAA J 47(2): 314–321

Große S, Schröder W (2009c) Wall-shear stress patterns of coherent structures in turbulent duct flow. J Fluid Mech 633: 147–158

Große S, Schröder W (2009d) High Reynolds number turbulent wind tunnel boundary layer wall-shear stress sensor. J Turbul 10(14): 1–12

Große S, Schröder W, Brücker C (2006) Nano-newton drag sensor based on flexible micro-pillars. Meas Sci Technol 17(10): 2689–2697

Große S, Soodt T, Schröder W (2008) Dynamic calibration technique for the micro-pillar pillar shear-stress sensor MPS³. Meas Sci Technol 19(10): 105 201

Humphrey JAC, Barth FG (2007) Medium flow-sensing hairs: biomechanics and models. In: Casas J, Simpson SJ (eds.) Advances in insect physiology, Academic Press 34: 1–80

Humphrey JAC, Devarakonda R, Iglesias I, Barth FG (1993) Dynamics of arthropod filiform hairs I. Mathematical modelling of the hair and air motions. Phil Trans R Soc Lond B 340: 423–444

Krijnen GJM, Dijkstra M, Baar JJ, Shankar SS, Kuipers WJ, Boer RJH, Altpeter D, Lammerink TSJ, Wiegerink R (2006) MEMS based hair flow-sensors as model systems for acoustic perception studies. Nanotech 17: 84–89

Krupa DJ, Matell MS, Brisben AJ, Oliveira LM, Nicolelis MAL (2001) Behavioral properties of the trigeminal somatosensory system in rats performing whisker-dependent tactile discriminations. J Neurosci 21: 5752–5763

Liu C (2007) Micromachined biomimetic artificial haircell sensors. Bioinspiration Biomimetics 2: 162–169

Ozaki Y, Ohyama T, Yasuda T, Shimoyama I (2000) An air flow sensor modeled on wind receptor hairs of insects. Proc MEMS, Miyazaki, Japan, 531–536

Padmanabhan A (1997) Silicon micromachined sensors and sensor arrays for shear-stress meas-

urements in aerodynamic flows. PhD thesis, Department of Mechanical Engineering, MIT, USA

Raffel M, Willert CE, Wereley ST, Kompenhans J (2007) Particle image velocimetry – A practical guide. Springer, Berlin Heidelberg New York

Shimozawa T, Kumagai T, Baba Y (1998) Structural scaling and functional design of the cercal wind-receptor hairs of cricket. J Comp Physiol A 183: 171–186

Suzuki Y, Kasagi N (1992) Evaluation of hot-wire measurements in wall shear turbulence using a direct numerical simulation database. Exp Thermal Fluid Sci 5: 69–77

Tucker C, Chen N, Engel J, Yang Y, Pandya S, Liu C (2006) High-sensitivity bi-directional flow sensor based on biological inspiration of animal haircell sensors. Proc 5th IEEE Sensors Conference, Daegu, Korea

van der Heiden K, Groenendijk BCW, Hierck BP, Hogers B, Koerten HK, Mommaas AAM, Groot ACG, Poelmann RE (2006) Monocilia on chicken embryonic endocardium in low shear stress areas. Dev Dyn 235: 19–28

Womersley J. R (1955) Method for the calculation of velocity, rate of flow and viscous drag in arteries when the pressure gradient is known. J Physiol 127: 553–563

Yang Y, Nguyen N, Chen N, Lockwood M, Tucker C, Hu H, Bleckmann H, Liu C, Jones DL (2010) Artificial lateral line with biomimetic neuromasts to emulate fish sensing. Bioinspiration Biomimetics 5: 106 001

Zou J, Chen J, Liu C (2001) Plastic deformation magnetic assembly (PDMA) of out-of-plane microstructures: Technology and application 3029. J MEMS 10: 302–309

Design and fabrication process for artificial lateral line sensors

28

Nima Izadi, Gijs J. M. Krijnen

Contents

Abstract

Biological sensory systems often display great performance, inspiring engineers to develop artificial counterparts. The lateral line system of fish has been widely studied by biologists for its crucial role in fish behaviour. Moreover recently the robustness, sensitivity and consequently wide range of applications that potentially benefit from the abilities of such a system have attracted the attention of the engineering community. Aquatic flow sensors based on the lateral line of fish are useful in underwater robotic applications for hydrodynamic imaging of complex and noisy environments to provide information for, for example, surveillance, navigation and obstacle detection. The speculative ability to manoeuvre in murky or dark water, especially object detection at short range, has strongly driven research in this area.

This chapter begins with a brief review of the state of the art flow sensors inspired by neuromasts. Subsequently, the design principles and fabrication scheme for a differential capacitive flow sensor incorporating a hair-like structure are discussed. The main focus is on the application of Micro-Electro-Mechanical Systems (MEMS) technology, which enables fabrication of sensory structures on length-scales comparable to what can be found in nature. Its intrinsic batch fabrication capability and high spatial resolution facilitate the construction of dense arrays of flow sensors, eventually allowing the function of the fish lateral line to be copied.

Nima Izadi
University of Twente
Transducer Science and Technology
P. O. Box 217, 7500 AE Enschede, The Netherlands
e-mail: n.izadi@ewi.utwente.nl

1. Introduction

Aquatic environments have been the cradle of life ever since its earliest emergence on the planet. Fish, the inhabitants of the vast

oceans, have struggled through 530 millions years of existence and adapted to their resourceful but challenging surroundings. The need to perceive and locate prey and to escape from predators, or in general to survive in highly competitive situations, has driven such living organisms to develop various sensory organs. Whereas more well-known sensory systems, such as the visual system, have played their roles in sustaining life, fish and amphibians are in addition gifted with a sensory organ called the lateral line.

The lateral line system enables fish to use the velocity profile of the surrounding water to construct a 3D map of their immediate environment on which depends their ability to school, localise prey or predators and to avoid obstacles (Dijkgraaf 1963, Coombs 2001). In some species like the blind cave fish, which lacks visual capabilities, fish survival predominantly depends upon this system (von Campenhausen et al. 1981).

The lateral line system consists of mechanoreceptive hair cells, covered by a jelly-like cupula, which together are referred to as neuromasts. Neuromasts are either located on top of the skin at the bottom of a visible pit or groove (superficial neuromasts) or in the lateral line canals along the body (canal neuromasts). The displacement of the cupula due to fluid motion couples to the stereovilli of the hair cells and changes the firing rate of the afferent neurons (Voigt et al. 2000). Using this system, fish are able to perceive cupula displacements as small as a few nanometers which corresponds to a few microns per second fluid velocity (van Netten 2006).

Man-made aquatic flow sensors have been built based on many sensing principles including heat conduction[1] (Chen and Liu 2003), the Doppler effect[2] (Jentink et al. 1987) and pressure difference[3] (Fernandez et al. 2007), but they are still rarely utilized for underwater applications. Small, robust, cheap and low energy consuming sensors are needed in marine environments for numerous applications (Arshad 2009). In recent years, however, benefiting from the advancement in technology and the knowledge and inspiration gained from nature, engineers have started to develop biomimetic flow-sensors aimed at a high level of versatility, robustness and sensitivity.

Biomimetics or bionics strives to exploit the principles of natural systems and to implement them in engineering products using available or new technologies. Lately, the need for more advanced and robust systems and the biological discoveries of functional properties of organic systems have popularized this perspective among engineers (Vincent 2001).

Lateral line inspired flow-sensor arrays can be used to study hydrodynamic phenomena in complex, noisy environments. They provide invaluable means for the study of fluid mechanics especially of turbulent flow and can also potentially be used in Underwater Autonomous Vehicles (UAV) to increase manoeuvre efficiency (Arshad 2009).

[1] Hot-wire anemometers measure fluid velocity by detecting temperature changes of heated wires due to heat convection forced by fluid flow. However, their function is limited at low flow velocities due to their relatively low sensitivity.

[2] Laser Doppler velocimeters point a monochromatic laser beam towards a target and collect the reflected radiation. Due to the Doppler effect, the change in wavelength of the reflected radiation is a function of the relative velocity of the targeted object. They are usually complicated devices and need reflective particles. Moreover, sensors based on acoustic transmission and detection of Doppler phase shift are generally large.

[3] Pressure distribution measurements can be used as an indirect way of flow sensing. However, information about the main stream is not accurately provided by this method.

2. Piezoresistive sensors

Piezoresistivity (Smith 1954) is a property of certain materials which change their electrical resistance when subject to tensile or compressive stresses. The effect is small in metal conductors but easily observable and exploitable in semiconductors such as silicon. It is a function of doping, temperature, crystallographic orientation, etc. Piezoresistive readout is commonly used to measure the deflection of members, such as cantilevers (see below) and membranes. It is possible to deduce the magnitude of force on a cantilever using piezoresistive regions (strain gauges) at the base, where the maximum stress occurs, by measuring resistance changes. These strain gauges are usually arranged in a Wheatstone bridge configuration and the change in resistance is proportional to the strain exerted on them. This relation is formulated as

$$\frac{\Delta R}{R} \propto \varepsilon \tag{1}$$

where R is the electrical resistance of the gauge, ΔR the change in R and ε the mechanical strain.

A *cantilever* is a beam that is fixed at one end and free at the other. Cantilevers at the micron-scale are made by micromachining, and have a wide range of applications in science and technology such as in AFM (Atomic Force Microscopy), chemical detection sensors, RF (Radio Frequency) filters and resonators, to name but a few. The basic principle of a cantilever-based sensor is to determine the magnitude of a load from the amount of deflection. It has been shown (Gere 2003) that for a cantilever subjected to a transverse load, the maximum strain is at the base (fixed end) and is equal to

$$\varepsilon_{\max} = \frac{M \cdot c_1}{E \cdot I} \tag{2}$$

in which M is the magnitude of the developed moment at the base, E is the Young's modulus of elasticity, I is the second moment of area and c_1 is the distance from the neutral axis. For cantilevers made of an isotropic and homogeneous material $c_1 = t/2$ in which t is the thickness of the beam.

2.1 Sensors based on cantilevers

Fan et al. (2002) have realized a cantilever-based flow sensor with a piezoresistive readout mechanism. A PDMA[4] process was utilized to obtain (820 µm long) out-of-plane cantilevers from in-plane fabricated cantilever beams. The fabrication starts with the selective doping of the silicon substrate with boron to form strain gauges. A backside etching process in alkaline (KOH) solution determines the cantilever's thickness. Thereafter, lead wires, a copper sacrificial layer and a thin gold layer, are deposited and patterned. Then Permalloy, which is necessary for the PDMA process, is electroplated. Subsequently, the cantilever is completed with an etching process on the front side. The sacrificial layer is removed in a wet etching process and, using a magnetic field, the Permalloy structure is raised to a permanent upright position. Finally a thin layer

[4] Plastic Deformation Magnetic Assembly (PDMA) is a technique for three-dimensional assembly of micro structures. Certain parts of the structure can plastically deform by applying external magnetic fields which interact with the magnetic material deposited on the micro structure. The resultant force brings the entire structure to a certain angle at which the developed stress can surpass the elastic limit of material and cause permanent deformation. PDMA is a non-contact, batch process but it is relatively difficult to accurately control the deformation angles of the structures.

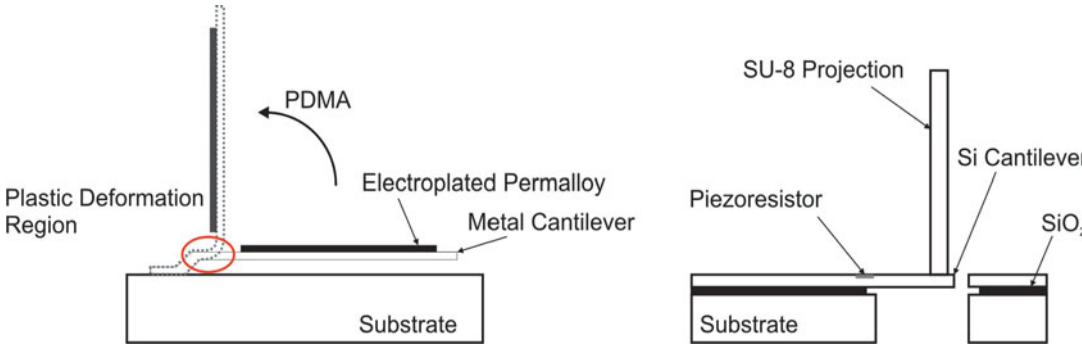


Fig. 1 **Left:** Schematic of cantilevers realised using PDMA (Fan et al. 2002, Chen et al. 2003). **Right:** Schematic of a SU-8 cylinder at the tip of a silicon cantilever (Yang et al. 2007)

of parylene is deposited in order to insulate the structure and the electrical conductors from water. The highest presented relative change of resistance, $\Delta R/R$, is 2.3×10^4 ppm at 1 m/s water flow speed.

The major advantage of this method is the monolithic fabrication process. However, according to the authors the robustness and compatibility of the device are a concern. The strength of the joint between out of plane and in plane cantilevers is crucial but it was shown that it is possible to strengthen the joint and define the exact bending point by electroplating a thin layer of gold.

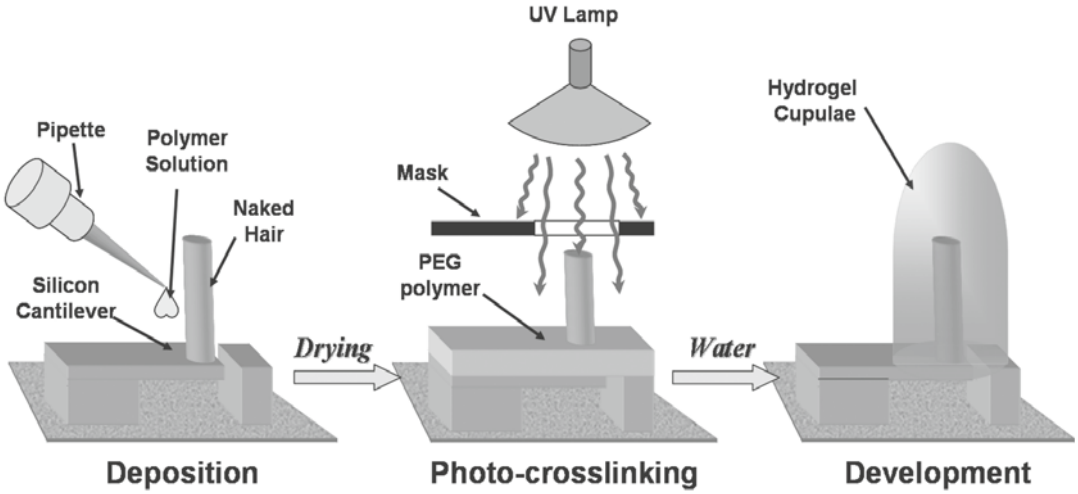
In another approach Chen et al. (2003) employed the same PDMA principle and produced a polymer-based flow sensor. Their fabrication consists of successive metal and polymer deposition and patterning. The main difference is in the use of *NiCr* strain gauges on the base of polyimide film which forms the out-of-plane cantilever (600 to 1,500 µm long). Low temperature processing (the highest process temperature reported is 350 °C) allows the use of a variety of substrates including flexible polymer materials. Chen et al. (2003) also have claimed that this process is more robust and efficient than using silicon bulk and surface micromachining. However, the strain gauge efficiency is considerably lower. The sensor has been tested in air flow and the highest reported relative change of resistance, $\Delta R/R$, is 600 ppm at 10 m/s air flow speed.

Yang et al. (2007) have fabricated a cantilever-based flow sensor using SU-8[5] polymer to produce a (500 µm long) cylindrical structure at the free end of a cantilever. The reported detection threshold[6] amplitude is 0.1 mm/s at 25 Hz using 2 Hz FFT (Fast Fourier Transform) bandwidth[7] (Chen et al., 2007) for measurements. The calculated resonance frequency is 1 kHz in water. The readout mechanism is the same as above, i. e. piezoresistive. The fabrication process begins with the use of Silicon On Isolator (SOI) wafers. Piezoresistive elements are then produced using ion implantation. Subsequently gold is

[5] SU-8 is a photo-sensitive polymer used in MEMS technology. It can be spun with different thicknesses and is used for the fabrication of high aspect ratio structures.

[6] Detection threshold is defined as the minimum detectable input. It depends on the properties of the sensor, the noise level and the measurement frequency resolution.

[7] Using a narrow FFT bandwidth (low measurement resolution) reduces the (Gaussian) noise power incurred in the measurement and, therefore, increases the signal to noise ratio which results in a lower detection threshold. However, it increases the detection time.



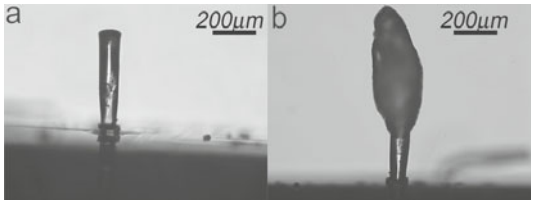


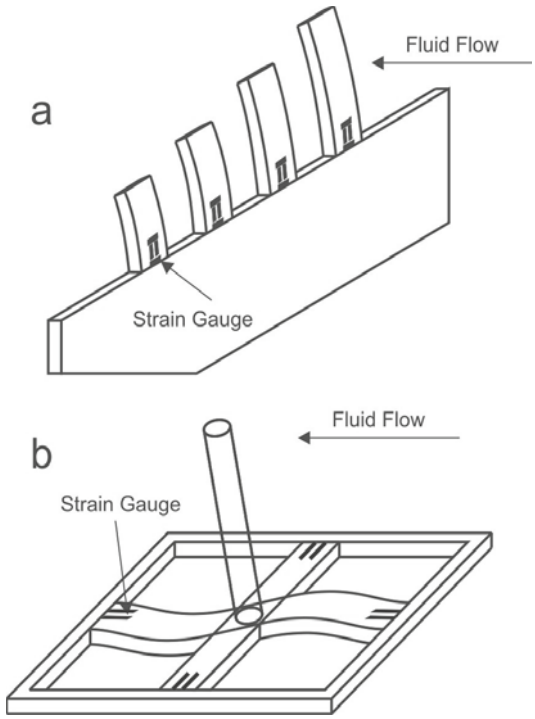
Fig. 2 **Top:** Schematic process steps to form a hydrogel cupula around the SU-8 hair at the tip of silicon cantilever (Peleshanko et al. 2007). Reproduced with permission of Wiley-VCH Verlag. **Right:** Microscopic pictures from a hair without **a** and with **b** hydrogel at the tip (McConny et al. 2009). Reproduced with permission of The Royal Society of Chemistry

deposited and patterned to form conductive wires and bond pads. This is followed by two DRIE[8] steps, both on the front and backside of the wafer, to define the cantilever. Subsequently, SU-8 photoresist is spun and patterned to shape the hair-like extension and then the cantilever is emersed in BHF (Buffered Hydro-Fluoric acid) solution. The devices have been made in pairs oriented perpendicular to each other to provide flow measurement along two orthogonal axes. These sensors have a linear response to AC flow and the hair-like SU-8 structures can be deflected by up to 35° without degradation of performance. Using this sensor, Yang et al. (2010) have produced a lateral line system capable of flow source localisation in a 3D domain.

Peleshanko et al. (2007) have reported a 20 to 70 fold increase in sensitivity[9] at frequencies between 10 to 110 Hz, and a decrease in the detection threshold from 100 µm/s to 75 µm/s water flow speed of the previously mentioned sensor. A water soluble Polyethylene Glycol (PEG) is dispensed on the SU-8 cylinder (the hair) and its immediate surroundings. Then UV-photopolymerization is carried out to cross link the polymer around the hair. After this, the sensor is put into water so that non-cross-linked polymer is dissolved and the cross-linked polymer swells and forms a dome-shaped artificial cupula around the hair. The increase in sensitivity has been reasoned to be a result of 1) a larger drag force due to

[8] Deep Reactive Ion Etching (DRIE) is a process to etch high aspect ratio structures in a silicon substrate.

[9] The sensitivity is defined as the ratio of a change at output to the respective change at input. It is basically the slope of the calibration curve of the sensor.



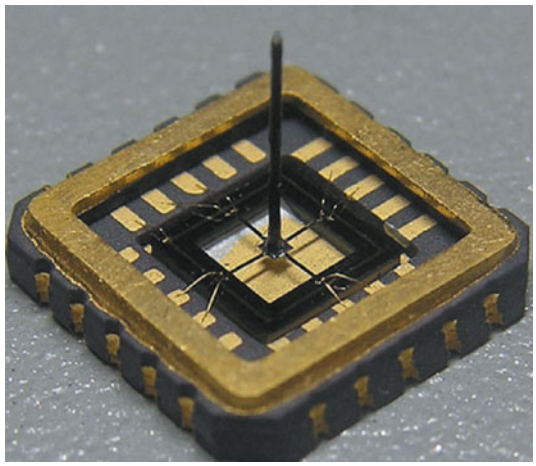


Fig. 3 **Left: a** In plane fabricated cantilevers are rotated to face the fluid flow. **b** A thin cylindrical wire is mounted at the joint of cross-shaped beams with piezoresistive elements at the base (Ozaki et al. 2000). **Top:** The hydrophone before packaging (Xue et al. 2007). Reproduced with permission of Elsevier

a larger effective cross section and 2) the coupling between the flow and the fluid trapped in the hydrogel (water content of the swallowed hydrogel is about 90 %). This approach also provides extra protection for the hair sensor. Using a similar principle, McConny et al. (2009) have reported a 38 fold increase in sensitivity and reduction of detection threshold to 2.5 µm/s. In this approach the dispense process was modified to increase the length of the hair and its cross section at the top without changing the mechanical characteristics of the base of the hair. Although the fabrication process is not monolithic and, therefore, is not suitable for array fabrication, according to the authors, it rivals the fish mechanoreceptors in terms of sensitivity (see chapter 23 by McConney and Tsukruk).

Another attempt to mimic hair receptors was made by Ozaki et al. (2000). They suggested two types of structure: 1) The first structure is rather simple, and is basically a series of planar cantilevers with piezoresistive strain gauges at the base. After fabrication the substrate is rotated by 90° to orient the cantilevers perpendicular to the direction of the fluid flow. The fabrication process is simple and very well developed in MEMS technology. The one dimensional nature of the fabricated array and its low density are two disadvantages of this approach. 2) The second configuration is a cross-shaped structure with a piezoresistive element at the end of each beam which is fixed to the substrate. A long metal wire is manually attached to the centre of the cross. This structure is made from thin (200 µm) silicon substrates which are first oxidized and patterned. Then Boron is diffused in order to make strain gauges. Anisotropic etching of the substrate from the backside in TMAH (TetraMethylAmmoniumHydroxide) is used to produce the beams. Afterwards aluminium interconnects are deposited and finally a Reactive Ion Etching (RIE) process from the front side is used to release the device. Although the fabrication process is simple, the

TMAH backside etching needs careful attention to accurately ensure the desired thickness. These structures have been made for and tested in air flow but potentially they can be used in water with minor modifications, mainly through the insulation of wires to prevent contact with water.

This second configuration has been repeated by Xue et al. (2007) and Zhang et al. (2008) who changed the piezo-transducer configuration and fabrication process and included external structures to adapt the sensor to aquatic environment. The hair is again attached manually. The sensor is connected to a 50 dB low noise preamplifier, immersed in Castor oil and protected using a Polychloroprene Rubber dome. The reported sensitivity is –197.7 dB (at 400 Hz) (0 dB = 1 V/µPa).

Lee et al. (2006) use a piezoelectric polymer film or PVDF (Polyvinylidene Fluoride), which is bonded to the surface of aluminium cantilever beams. The overall size of the device in this approach is bigger (cantilevers are between 22 to 30 mm long) than the previous ones as macro-fabrication technology has been utilized. Therefore, their sensitivity and bandwidth are much lower compared to micro-fabricated counterparts.

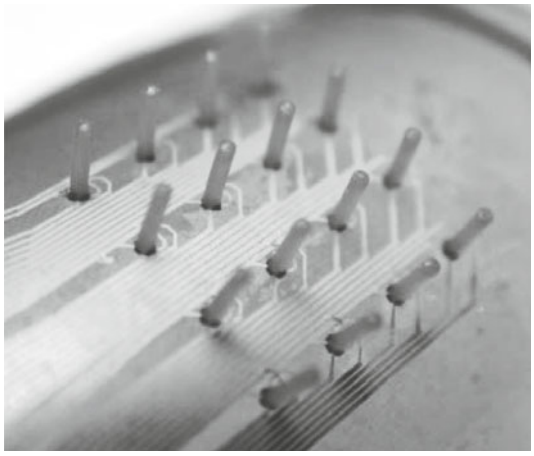


Fig. 4 Polyurethane hairs on flexible substrate (Engel et al. 2006). Copyright 2006 IEEE; reproduced with permission

2.2 All polymer sensors

Engel et al. (2006) have reported another type of hair-based flow sensor using Polyurethane elastomers and Force Sensitive Resistors (FSR). This consists of a hair-like structure on top of four FSRs which are in half bridge configuration and can reveal both magnitude and direction of displacement. When the structure deflects, the stress develops at its base and the resistance of the FSRs will change accordingly. The building process begins with the deposition and patterning of gold lead wires. Afterwards a thick photoresist is spun and patterned to form a mould for the FSRs. The actual FSRs are made of polyurethane that is loaded with an electrical conductive filler. This filler can be Carbon Black (CB) or multiwall carbon nanotubes (MW-CNT). The FSR material is applied to the substrate and sacrificial photo resist is removed. The hairs are made by filling a wax mould with the polymer and are then aligned and attached to the FSR substrate. Subsequently, the wax mould is dissolved in hot water. These structures exhibit a great robustness but suffer from non-uniformity, viscoelastic creep, size and aspect ratio limitations (because the wax mould is formed by drilling) and cross-axis coupling (Engel et al. 2006). The latter can be partially overcome using optimized cross-sections.

2.3 Conclusion

The above mentioned sensors all use the piezoresistive effect as the readout mechanism. Piezoresistive readout has the advantage of being cheap and easy to fabricate. Moreover, the readout circuitry can be rather simple. The development of these sensors has already led to the production of artificial lateral line sensor arrays. However, piezore-

sistors are sensitive to temperature changes while the magnitude of change in resistance due to the stress is comparatively small with respect to the base resistance, i. e. the signal modulation is small. In addition, to decrease the thermal noise the base resistance needs to be small. When the base resistance decreases, the electrical current supply should increase to yield a strong signal, hence the power consumption increases.

3. A capacitive sensor

A capacitor consists of a pair of electrical conductors separated by an insulating material (dielectric). When a voltage difference is applied to the conductors an electric field is established in the dielectric material. Energy is stored in this electric field. The ability of a capacitor to store energy is characterised by its capacitance. An electrical capacitance C is defined as the ratio of the stored electric charge Q induced by establishing the electric potential V between the two electrodes. When these electrodes are parallel plate conductors it can be shown that[10]

$$C = \frac{\varepsilon \cdot A}{d} \tag{3}$$

is a physical characteristic of the configuration in which A is the overlapping area of the electrodes, d is the distance between them and ε is the electrical permittivity of the medium in between. A change of any of these three quantities is reflected in the magnitude of the capacitance and can be measured with suitable electrical circuits. A capacitive sensor is a transducer which converts a stimulus to a corresponding change in a certain capacitance. A comprehensive study on the capacitive measurement principle can be found in de Jong (1994) and Baxter (1997).

A basic requirement for using capacitive readout in an aquatic environment is to prevent water from coming into contact with the electrodes. This may result in a short circuit or electrolysis if the water has sufficient conductivity. To overcome this difficulty conductors used in aquatic environments have to be insulated. In MEMS technology an option for achieving this is to deposit or spin a layer of a non-conductive polymer over the chip as mentioned in some of the previously noted devices. In addition, a capacitive change is usually due to a change in the distance of the two electrodes. Therefore the presence of a highly viscose medium in between the electrodes can degrade the performance of the device considerably[11] (Blech 1983).

3.1 Analysis and design

Figure 5 shows the concept of a capacitive aquatic hair-based flow sensor. A high aspect ratio[12] structure (the hair) is firmly attached to a fully supported membrane. The readout mechanism is beneath the membrane and is insulated from the fluid flow. This consists of two electrodes attached to the bottom side of the membrane and a common electrode at a distance below them. The configuration forms two capacitors, which can be used in a differential readout scheme reducing disturbing common mode signals such as

[10] This is the case when the effect of the fringing electrical field is neglected, i.e. when the lateral dimensions of the electrodes are much larger than their separation gap.

[11] This effect is called squeeze film damping.

[12] Aspect ratio in this case is the ratio of the length of hair to its diameter.

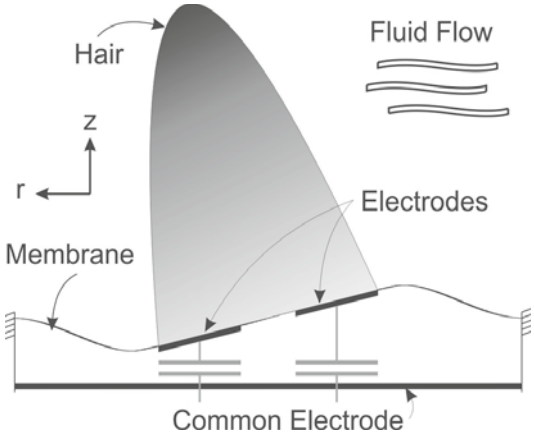


Fig. 5 Schematic view of a capacitive aquatic flow sensor. For clarity, the capacitances are shown by electrical symbols

changes in temperature, humidity, etc. Having the hair on a membrane forming a cavity devoid of water lessens the chances of short circuiting, prevents electrolysis and reduces squeeze film damping effects as discussed previously.

The hair deflects as a result of the drag force exerted by the fluid flow. This deflection is mirrored in the flexible membrane to which the hair is firmly attached. The distance between each of the electrodes and the common one, or in other words, the gap between the capacitor plates is then altered, as one gap increases while the other one decreases. By measuring the resulting change in capacitances, the amount of membrane and hair deflection, and hence the magnitude of the drag-force, thus the flow-velocity of the water can be deduced.

Geometrical optimization that takes into account technological possibilities and limitations is needed to ensure the realisation of a flow-sensor with the desired characteristics. To this end, a relation between the capacitance change and the flow velocity needs to be derived. This problem is divided into three parts and discussed in what follows.

3.1.1 On the magnitude of the force on the hair structure due to a specific flow velocity

The relation between an oscillatory fluid flow and the drag it exerts on a rotatable cylinder has been thoroughly dealt with in the literature, for example, in the context of the research on the principles of deflection of filiform hairs on cricket cerci and spider legs (see e.g. Shimozawa et al., 1998 and Humphrey and Barth, 2007). In brief, the fluid flow velocity profile close to the surface is determined first. Next, the hair structure is modelled as a rigid cylinder in a laminar flow and the drag force it incurs is calculated. To this end, the flow regime should be determined considering Reynolds and Strouhal numbers (Humphrey et al. 1993). Eventually, using a second order system approximation, the equation of motion is written as

$$I\frac{\partial^2\theta_h}{\partial t^2}+R\frac{\partial\theta_h}{\partial t}+K_\varphi\theta_h=\int_0^L F_s(z,t)\cdot z\cdot dz \quad (4)$$

in which L is the length of the hair, z is the distance from the substrate, F_s is the drag force and θ_h is the steady state response (the angle of deflection) of the system. The moment of inertia I is a result of the geometry of the structure with additional contributions due to fluid cylinder interaction. The damping coefficient R consists of the intrinsic structural damping and viscous damping of the fluid. In our sensor geometry, squeeze film damping needs to be added to this term. The torsional resistance K_φ depends on the mechanical characteristics of the structure and can be derived analytically.

3.1.2 On the deflection of the structure due to a specific torque

The right hand side of (4) is the torque that is generated by the drag force on the hair. It can be shown (Young and Budynas 2002) that for a fully supported plate with a trun-

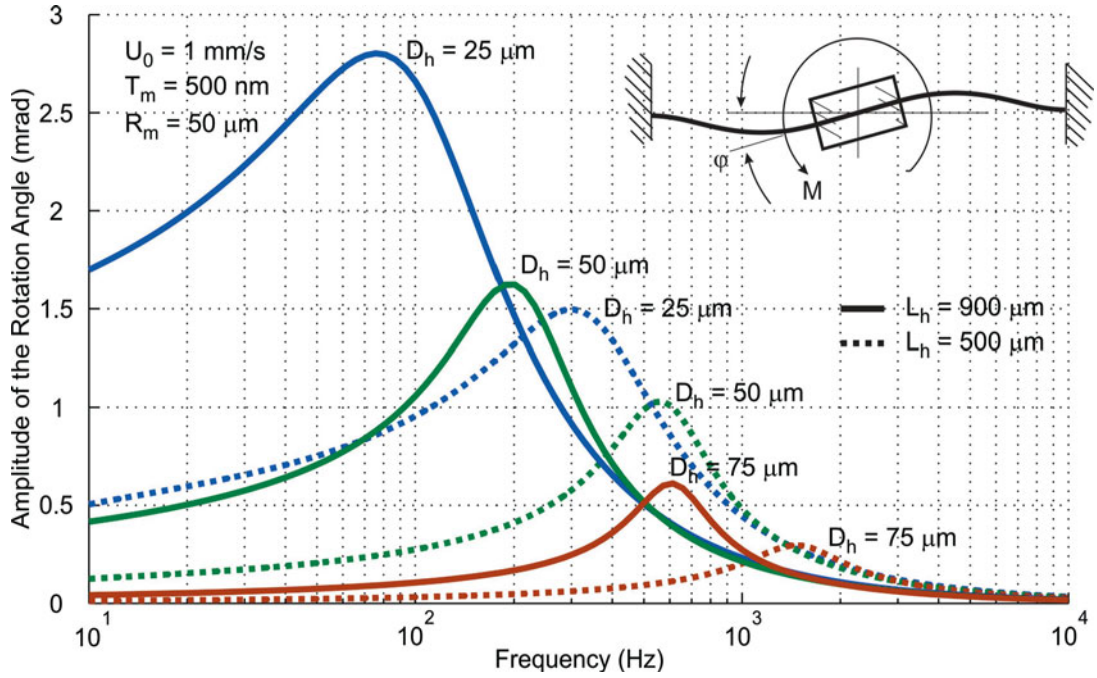


Fig. 6 Rotation angle as a function of frequency for different hair lengths and diameters. The inset shows the schematic view of the anti-symmetric deflection of a fully supported membrane

nion at the centre (Fig. 6, inset) the angle of deflection is related to the applied torque as

$$\varphi = \frac{\alpha(\zeta,\nu)\cdot M}{E\cdot T_m^3} \tag{5}$$

in which E is the Young's Modulus of elasticity, M is the torque acting on the hair, T_m is the membrane thickness and $\alpha(\zeta,\nu)$ is a constant which depends on Poisson's ratio ν and the ratio of the diameters of the trunnion (R_h) and the membrane (R_m). For $\zeta = R_h/R_m = 0.5$ and $\nu = 0.3$, it is derived that $\alpha = 0.081$ (Timoshenko and Woinowsky-Krieger 1959). It is intuitively clear that for a specific thickness of the membrane the flexibility increases as the ζ ratio decreases. However, the fabrication of thin hairs and large membranes has some limitations. In addition, the drag force on a thin hair is relatively small. Figure 6 shows the deflection of a 500 nm thin SU-8 membrane.

The stiffness of the membrane is also affected by its shape. The above calculations have been made based on the use of a circular membrane. However, it can be shown that an elliptical membrane with its major axis along the desired direction of measurement shows a larger deflection than a circular one for given torque. For these membranes an additional design parameter that should be taken into consideration is the ratio of the semi-major to semi-minor axes.

3.1.3 On the relative capacitance change due to a specific deflection angle of the hair structure

The capacitance depends on the area and the shape of electrodes. At the initial state (no deflection) the capacitance is easily obtained using (3). To calculate the capacitance while the membrane is deflected, C_f, we need to evaluate

$$C_f = \int\int \frac{\varepsilon}{d_0 + w(r,\theta)}\cdot r\cdot dr\cdot d\theta \tag{6}$$

Table 1 The design parameters and resulting characteristics of the sensor

Material (SU-8)		**Material (Water)**	
Density	1 200 kg/m^3	Density	1 000 kg/m^3
Young's Modulus (E)	4.4 GPa	Dynamic Viscosity	1.002×10^{-3} Pa · s
Poisson's Ratio	0.22	Speed of Sound	1 500 m/s
Geometrical Dimensions			
Radius of the membrane (R_m)	50 µm	Length of the hair (L_h)	~900 µm
Thickness of the membrane (T_m)	~500 nm	Initial gap between	
Radius of the hair (R_h)	~25 µm	electrodes (d_0)	~1 µm
Characteristics (calculated)			
Resonance Frequency	~200 Hz	Detection Threshold	~2.2 mm/s at 100 Hz
Quality Factor	~1	Sensitivity	~154 fF/mm/s at 100 Hz

in which r and θ denote the polar coordinates, d_0 is the initial gap between the electrodes and $w(r,\theta)$ is the analytical solution for the deflection of the membrane. This integral needs to be solved numerically. The more exact measure can be obtained using Finite Element Analysis (FEA). Table 1 above summarises the design parameters, geometrical dimensions and the resulting characteristics of the sensor.

3.2 Fabrication

The fabrication process has been thoroughly described in Izadi et al. (2010) and is schematically shown in Fig. 7. It starts with a highly doped silicon wafer (which acts as a common electrode). At the first step a DRIE process is used to make 60 µm deep and 4 µm wide trenches (etch ports) on the front side of the wafer. This is followed by an oxidation process and the deposition of a stoichiometric silicon nitride (Si_3N_4) layer. Subsequently, polysilicon is deposited to close the

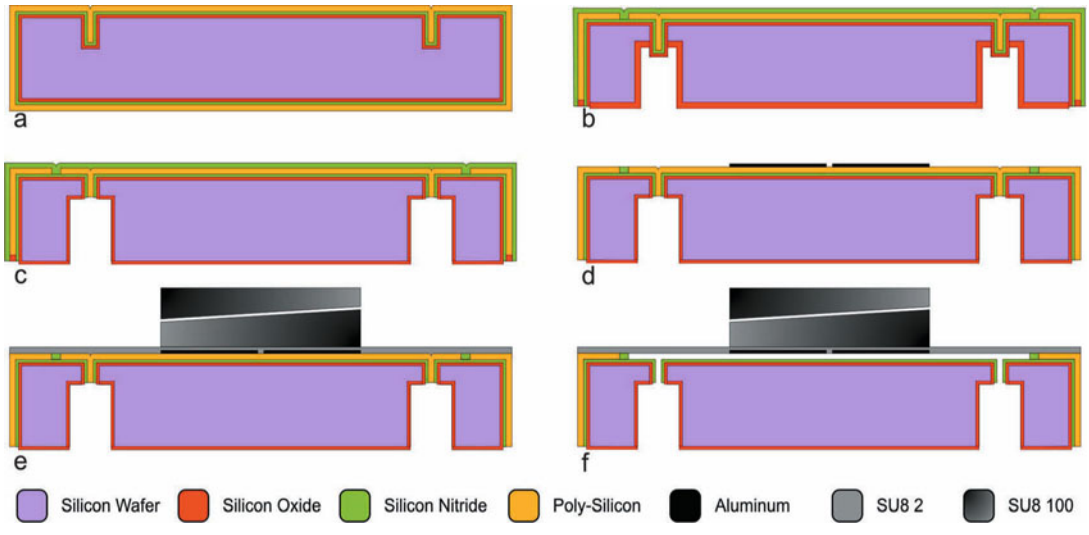
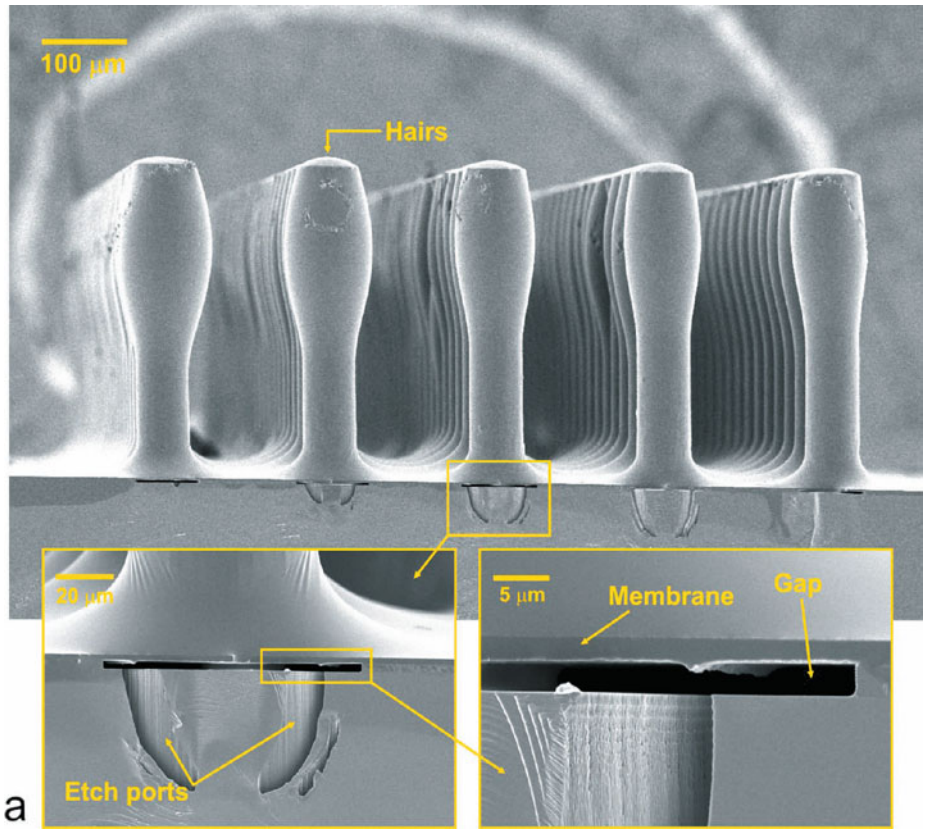


Fig. 7 Abridged fabrication process



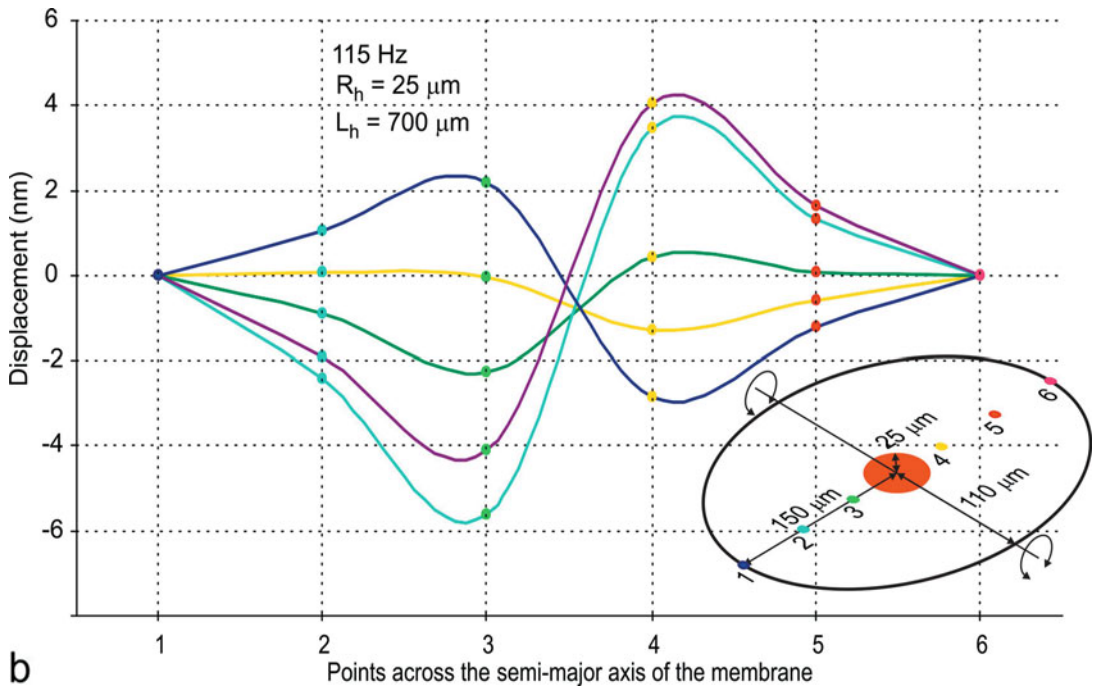


Fig. 8 **a** SEM picture of capacitive hair based flow sensor. **b** Displacement of an elliptical membrane measured across its semi-major axis with laser vibrometer in air

etch port openings. Therefore, the minimum thickness of the polysilicon layer is determined by the width of the etch ports, and the thickness determines the gap between the electrodes. To define membranes, a rim is etched and re-filled with Silicon Rich silicon Nitride (SiRN). After a back-side DRIE step and wet oxidation, a directional SF_6 based plasma with enhanced ion bombardment is used to open the protection stack at the etch ports and expose the sacrificial polysilicon layer. Then, the SiRN layer is removed from the front side and low stress aluminium electrodes are deposited. A thin layer (~500 nm) of SU-8 is spun and selectively exposed to form the membranes. This is followed by a thick SU-8 (500 to 800 µm) spin, exposure and development step to form the hairs. The shape of the hairs, using the negative tone properties of SU-8, is tuned to increase the drag force on them by increasing the diameter at the top. Finally, using a selective SF_6 based high density plasma without self bias, the sacrificial layer is etched to release the membranes (see Fig. 8a).

Using a laser vibrometer, preliminary mechanical tests of the fabricated sensor have been made in air. At the right side of Fig. 8 the displacement of an elliptical membrane in the half cycle of oscillation at different phases is shown. It should be noted that the small asymmetric deflections of the membranes are a result of a slight misalignment of the hair relative to the membrane.

3.3 Distributed sensory array

Although a single sensor may be capable of fluid flow detection, the full functionality of the lateral line systems relies on the interpretation of parallel data streams from an array of sensors (Casas et al. 2010). Such an array provides both spatially and temporally rich data that can be used to construct a 3D map of the immediate environment. Fortunately, MEMS technology enables us to fabricate arrays, as can be seen in Fig. 8. The fabrication of hairs with different diameters can be carried out in a single lithography step. Various hair lengths can be achieved using multiple SU-8 patterning processes. However, as the number of layers increases the processing becomes more difficult. The differences in diameter and length translate directly into variations in frequency characteristics. The result is that the sensor dynamic range and frequency response are widened.

The distance of the hairs along the stream of flow is another important design parameter. The boundary layer and wake beyond a hair can have a substantial effect on the drag force on the next hair. This phenomenon is called viscous coupling and theories have been developed to predict and control it (Bathellier et al. 2005, Cummins et al. 2007).

3.4 Discussion

The performance of the sensor has at least two dimensions: the sensitivity and the bandwidth. Evaluation of the collective influence of geometrical dimensions and material properties on these characteristics of the sensor highlights the decisive parameters in the performance. These parameters are the thickness of the membrane, the ratio of the hair diameter to that of the membrane, denoted by ζ, the length and diameter of the hair and the Young's modulus of the membrane material.

Clearly, the membrane should be made of a compliant material with low Young's modulus. SU-8 is compatible with micro-fabrication technology, has a comparatively low Young's modulus among the typical "MEMS" materials, can be spun to form very thin or very thick layers and is a photo structurable polymer, hence a reasonable choice for both membrane and hair. To obtain a lower tor-

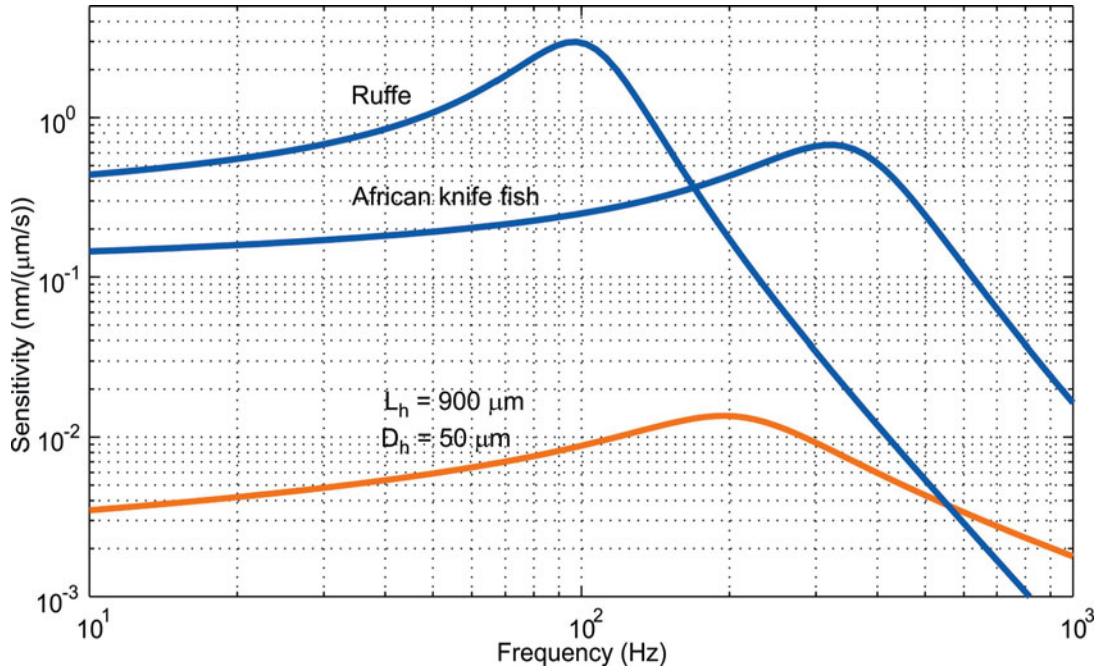


Fig. 9 Comparison of the mechanical sensitivities of the capacitive sensor and fish neuromast

sional stiffness, ζ should be small, thus the membrane area should be large and the hair should be thin. However, a membrane with a large area is more difficult to produce and is more susceptible to stiction. At the same time, a thin hair experiences less drag. Therefore, these parameters should be optimised to yield the highest sensitivity for a desirable bandwidth. On the other hand, the aspect ratio of the hair, i. e. length to diameter ratio, is technologically limited; the thinner the hair, the shorter it is. Therefore, the length of the hair, which should be maximised, is also linked to its diameter. Note that a thin, long hair has advantage over a short, thick hair even when ζ is constant. It can be shown that when the membrane radius is 50 µm, considering an aspect ratio of 20, the highest sensitivity is achieved with a hair diameter between 20 to 25 µm.

The thickness of the membrane has the greatest effect on the performance of the sensor and should be minimised. Using SU-8, the minimum achievable thickness of a defect free membrane is around 400 nm. The use of other materials, for example, Parylene, which can be deposited in even thinner layers, e. g. 100 nm, potentially leads to a more sensitive sensor.

In short, to achieve higher sensitivity the hair should be long, the membrane thin and flexible and ζ should be small. The excessive downward deflection due to the static hydrostatic pressure across a thin membrane can be avoided by adjusting the pressure of the back chamber. In the frequency and amplitude range of interest, the resultant increase of the squeeze film damping due to the higher pressure does not affect the dynamic behaviour of the system.

The mechanical sensitivity of the sensor, defined as the mean displacement of one electrode per unit flow velocity, is shown in Fig. 9 and compared with that of fish[13] (van

[13] The sensitivity of fish has been defined as the neuromast displacement per unit flow velocity.

Netten 2006). It is seen that using the proposed geometry and materials, the sensitivity of the sensor is around two orders of magnitude below its natural counterpart. In order to achieve a higher sensitivity, as discussed above, the drag pickup mechanism should be further improved and a thinner, larger and hence more flexible membrane should be employed. Furthermore, the overall performance is affected by the electrical sensitivity which is determined by the parasitic capacitances and the minimum detectable relative change of capacitance, $\Delta C/C$, both of which should be reduced by improvement of the readout circuit or possibly by integration of it with the sensor.

3.5 Conclusion

Inspired by the function of the neuromast organ of the lateral line system, a hair-based aquatic flow velocity sensor and its fabrication process have been proposed. Capacitive sensing is chosen as the readout mechanism as it offers accuracy, high resolution and low power consumption. However, several difficulties arise from using the capacitive sensing principle in a conductive medium with high density and viscosity, like water. These stem mainly from the electrical insulation of the electrodes (to prevent electrolysis and short circuit), added inertia of the movable electrodes and squeeze film damping. The result is a rather complicated, but feasible, fabrication process.

Summary

The research covering different sensory organs of animals is a rich field. Evolutionary pressure forces virtually each part of a living creature to adapt to the organisms' environmental conditions to ensure survival and reproduction. Biological studies of sensory organs reveal solutions emerging from nature that provide vital information for an animal. Recently attention and effort of engineers has been drawn to mimic nature's blueprints and to fabricate artificial counterparts of these sensory organs with the aim of reaching a higher level of performance and robustness necessary in modern technological applications.

The lateral line of fish has received great attention in recent years, and we have briefly reviewed here some of the efforts to mimic this system and to detail the design and fabrication of a capacitive sensor approach which promises great sensitivity and low power consumption.

The most powerful feature of the lateral line is its property as an array. This enables fish to construct a 3D view of their environment, rather than just having a means for velocity measurement, and enables them to navigate in murky water or even if they are blind. MEMS technology provides us with the ability to fabricate large sensor arrays in monolithic processes. These artificial sensors will benefit the study of fluid dynamics and the study of aquatic animal behaviour, and are key to improving the performance of underwater robotic systems.

References

Arshad MR (2009) Recent advancement in sensor technology for underwater applications. Indian J Mar Sci 38: 267–273

Bathellier B, Barth FG, Albert JT, Humphrey JAC (2005) Viscosity-mediated motion coupling between pairs of trichobothria on the leg of the spider *Cupiennius salei*. J Comp Physiol A 191: 733–746

Baxter LK (1997) Capacitive sensors: design and application. IEEE Press series on Electronics Technology

Blech JJ (1983) On isothermal squeeze films. J Lubr Technol 105: 615–620

Casas J, Steinmann T, Krijnen G (2010) Why do insects have such a high density of flow-sensing hairs? Insights from the hydromechanics of biomimetic MEMS sensors. J R Soc Interface, doi: 10.1098/rsif.2.010.0093

Chen J, Engel JM, Liu C (2003) Development of polymer-based haircell using surface micromachining and 3D assembly. In: Proceedings of the 12th International Conference on Solid State Sensors, Actuators and Microsystems, Transducer 2003, Boston, pp 1035–1038

Chen J, Liu C (2003) Development and characteri-

zation of surface micromachined out-of-plane hot-wire anemometer. IEEE ASME J Microelectromech Syst 12: 979–988

Chen N, Tucker C, Engel JM, Yang YC, Pandya S, Liu C (2007) Design and characterization of artificial haircell sensor for flow sensing with ultrahigh velocity and angular sensitivity. IEEE ASME J Microelectromech Syst 16: 999–1014.

Coombs S (2001) Smart skins: information processing by lateral line flow sensors. Autonomous Robots 11: 255–261

Cummins B, Gedeon T, Klapper I, Cortez R (2007) Interaction between arthropod filiform hairs in a fluid environment. J Theor Biol 247: 266–280

de Jong JW (1994) Smart capacitive sensors. Delft University press, Delft

Dijkgraaf S (1963) The functioning and significance of the lateral-line organs. Biol Rev 38: 51–105

Engel JM, Chen J, Liu C, Bullen D (2006) Polyurethane rubber all-polymer artificial hair cell sensor. IEEE ASME J Microelectromech Syst 15: 729–736

Fan Z, Chen J, Zou J, Bullen D, Liu C, Delcomyn F (2002) Design and fabrication of artificial lateral line flow sensors. J Micromech Microeng 12: 655–661

Fernandez VI, Hou SM, Hover FS, Lang JH, Triantafyllou MS (2007) Lateral line inspired MEMS-array pressure sensing for passive underwater navigation. In: Proceedings of the International Symposium on Unmanned Untethered Submersible Technology, UUST 2007, Autonomous Undersea Systems Institute

Gere JM (2003) Mechanics of materials, 6th edn. Thompson-Engineering

Humphrey JAC, Barth FG (2008) Medium flow-sensing hairs: biomechanics and models. In: Casas J, Simpson SJ (eds) Advances in insect physiology. Insect mechanics and control, Vol. 34, Elsevier Ltd, pp 1–80

Humphrey JAC, Devarakonda R, Iglesias I, Barth FG (1993) Dynamics of arthropod filiform hairs. I. Mathematical modelling of the hair and air motions. Phil Trans R Soc Lond B 340: 423–444

Izadi N, de Boer MJ, Berenschot JW, Krijnen GJM (2010) Fabrication of superficial neuromast inspired capacitive flow sensors. J Micromech Microeng 20: 085 041

Jentink HW, van Beurden JAJ, Helsdingen MA, Mul FFM, Suichies HE, Aarnoudse JG, Greve J (1987) A compact differential laser Doppler velocimeter using a semiconductor laser. J Phys E: Sci Instrum 20: 1281–1283

Krijnen GJM, Floris J, Dijkstra MA, Lammerink TSJ, Wiegerink RJ (2007) Biomimetic micromechanical adaptive flow-sensor arrays. In: Proc SPIE Europe Microtechnologies for the New Millennium, Maspalomas, Gran Canaria, Spain, pp 6592–6608

Lee YS, Im HJ, Kwon J, Yoon DJ (2006) Biologically inspired smart sensor for acoustic emission detection. Key Eng Mater 321–323: 204–207

Leydig F (1850) Über die Schleimkanäle der Knochenfische. Arch Anat Physiol Wiss Med: 170–181

McConney ME, Chen N, Lu D, Hu HA, Coombs S, Liu C, Tsukruk VV (2009) Biologically inspired design of hydrogel-capped hair sensors for enhanced underwater flow detection. Soft Matter 5: 292–295

Ozaki Y, Ohyama T, Yasuda T, Shimoyama I (2000) Air flow sensor modelled on wind receptor hairs of insects. In: Proc IEEE 13th International Conference on Micro Electro Mechanical Systems, MEMS 2000, Miyazaki, Japan, pp 531–536

Peleshanko S, Julian MD, Ornatska M, McConney ME, LeMieux MC, Chen N, Tucker C, Yang Y, Liu C, Humphrey JAC, Tsukruk VV (2007) Hydrogel-encapsulated microfabricated haircells mimicking fish cupula neuromast. Adv Mater 19: 2903–2909

Shimozawa T, Kumagai T, Baba Y (1998) Structural scaling and functional design of the cercal wind-receptor hairs of cricket. J Comp Physiol A 183: 171–186

Smith CS (1954) Piezoresistance effect in germanium and silicon. Phys Rev 94: 42–49

Timoshenko S, Woinowsky-Krieger S (1959) Theory of plates and shells, 2nd edn. McGraw-Hill

van Netten SM (2006) Hydrodynamic detection by cupulae in a lateral line canal: functional relations between physics and physiology. Biol Cybern 94: 67–85

Vincent JFV (2001) Stealing ideas from nature. In: Pellegrino S (ed) Deployable structures. Springer, Vienna, pp 51–58

Voigt R, Carton AG, Montgomery JC (2000) Responses of anterior lateral line afferent neurons to water flow. J Exp Biol 203: 2495–2502

von Campenhausen C, Riess I, Weissert R (1981) Detection of stationary objects by the blind cave fish *Anoptichthys jordani* (Characidae). J Comp Physiol A 143: 369–374

Xue C, Chen S, Zhang W, Zhang B, Zhang G, Qiao H (2007) Design, fabrication and preliminary char-

acterization of a novel MEMS bionic vector hydrophone. Microelectron J 38: 1021–1026

Yang Y, Chen N, Tucker C, Engel JM, Pandya S, Liu C (2007) From artificial hair cell sensor to artificial lateral line system: development and application. In: Proc IEEE 20th International Conference on Micro Electro Mechanical Systems, MEMS 2007, Kobe, Japan, pp 577–580

Yang Y, Nguyen N, Chen N, Lockwood M, Tucker C, Hu H, Bleckmann H, Liu C, Jones DL (2010) Artificial lateral line with biomimetic neuromasts to emulate fish sensing. Bioinspir Biomim 5: 016 001

Young WC, Budynas RG (2002) Roark's formulas for stress and strain, 7th edn. McGraw-Hill Education-Europe

Zhang B, Qiao H, Chen S, Liu J, Zhang W, Xiong J, Xue C, Zhang G (2008) Modelling and characterization of a micromachined artificial hair cell vector hydrophone. Microsyst Technol 14: 821–828

Index

N

O

P

R

S

T

V

W

List of contributors

Zane Aldworth
University of Washington
Department of Biology
Seattle, WA 98195-1800, USA
zane@uw.edu

Kentaro Arikawa
Graduate University for Advanced Studies
(Sokendai)
Laboratory of Neuroethology
Shonan Village
Hayama, Kanagawa 240-0193, Japan
arikawa@soken.ac.jp

Friedrich G. Barth
University of Vienna
Faculty of Life Sciences
Department of Neurobiology
Althanstr. 14
1090 Vienna, Austria
friedrich.g.barth@univie.ac.at

Daniel Bland
University of Queensland
Queensland Brain Institute
St. Lucia, QLD 4072, Australia

Horst Bleckmann
University of Bonn
Institute of Zoology
Poppelsdorfer Schloß
53115 Bonn, Germany
bleckmann@uni-bonn.de

Herbert Bousack
Forschungszentrum Jülich GmbH
Institute of Complex Systems
Peter Grünberg Institute
Bioelectronic (ICS-8/GPI-8)
52425 Jülich, Germany
h.bousack@fz-juelich.de

Bree Cummins
Tulane University
Center for Computational Science
Stanley Thomas Hall, Room 401
New Orleans, LA 70118, USA
bcummins@tulane.edu

Thomas Daniel
University of Washington
Department of Biology
Seattle, WA 98195-1800, USA
danielt@u.washington.edu

Fons De Mey
University of Antwerpen
Department of Mathematics and
Computer Science
Active Perception Lab
Prinsstraat 13
2000 Antwerpen, Belgium
Fons.Demey@ua.ac.be

Tobi Delbruck
University of Zürich and ETH Zürich
Institute of Neuroinformatics
Winterthurerstr. 190
8057 Zürich, Switzerland
tobi@ini.phys.ethz.ch

Jessica Fox
University of Washington
Department of Biology
Program on Neurobiology and Behavior
Seattle, WA 98195-1800, USA
jessfox@uw.edu

Tomas Gedeon
Montana State University
Department of Mathematics
Wilson Hall
Bozeman, MT 59715, USA
gedeon@math.montana.edu

Gregory J. Gerling
University of Virginia
Department of Systems and
Information Engineering
151 Engineer's Way
Charlottesville, VA 22904
gregory-gerling@virginia.edu

Sebastian Große
RWTH Aachen
Institute of Aerodynamics and Chair of
Fluid Mechanics
Wüllnerstr. 5a
52062 Aachen, Germany
sebastian.grosse@rwth-aachen.de

Hyuk Rok Gwon
Chung-Ang University
Heuksuk-Dong 221, Dongjak-Gu
Seoul 156-756, Korea
hrgwon99@gmail.com

Hossein Haj-Hariri
University of Virginia
Department of Mechanical and
Aerospace Engineering
122 Engineer's Way
Charlottesville, VA 22904-4746, USA
hh2b@virginia.edu

Mitra J. Z. Hartmann
Northwestern University
Departments of Mechanical Engineering
and Biomedical Engineering
2145 Sheridan Road
Evanston, IL 60208-3107, USA
m-hartmann@northwestern.edu

Armin Hinterwirth
University of Washington
Department of Biology
24 Kincaid Hall
Seattle, WA 98195-1800, USA
ahinterw@u.washington.edu

Tetsutaro Hiraguchi
Kobe University
Department of Biology
Rokkodai
Nada-ku
Kobe, Hyogo 657-8501, Japan

J. Sean Humbert
University of Maryland
Department of Aerospace Engineering
3181 Glenn L Martin Hall
College Park, MD 20742, USA
humbert@umd.edu

Joseph A. C. Humphrey†
University of Virginia
Department of Mechanical and Aerospace
Engineering
122 Engineer's Way
Charlottesville, VA 22904-4746, USA

Andrew M. Hyslop
University of Maryland
Autonomous Vehicle Laboratory
College Park, MD, USA
ahyslop@umd.edu

Nima Izadi
University of Twente
Transducer Science and Technology
P. O. Box 217
7500 AE Enschede, The Netherlands
n.izadi@ewi.utwente.nl

Masashi Kawasaki
University of Virginia
Department of Biology
Gilmer Hall
McCormick Road
Charlottesville, VA 22904, USA
mk3u@virginia.edu

Midori Kidokoro-Kobayashi
Kobe University
Department of Biology
Rokkodai
Nada-ku
Kobe, Hyogo 657-8501, Japan
midori@sapphire.kobe-u.ac.jp

Elmer K. Kim
University of Virginia
Department of Systems and Information Engineering
151 Engineer's Way
Charlottesville, VA 22904, USA
ek2f@virginia.edu

Adrian Klein
Universität Bonn
Institut für Zoologie
Poppelsdorfer Schloß
53115 Bonn, Germany
adrian@uni-bonn.de

Holger G. Krapp
Imperial College London
Department of Bioengineering
Royal School of Mines
South Kensington Campus
London, SW7 2AZ, UK
h.g.krapp@imperial.ac.uk

Gijs J. M. Krijnen
University of Twente
Transducer Science and Technology
P. O. Box 217
7500 AE Enschede, The Netherlands
gijs.krijnen@utwente.nl

Rafael Kurtz
Bielefeld University
Department of Neurobiology
Post Box 100131
33501 Bielefeld, Germany
rafael.kurtz@uni-bielefeld.de

Seong Hyuk Lee
Chung-Ang University
4331, Department of Mechanical Engineering
Heuk-Suk Dong
Dong-Jak Gu
Seoul, Korea
shlee89@cau.ac.kr

Fritz-Olaf Lehmann
University of Ulm
Institute of Neurobiology
Albert-Einstein-Allee 11
89081 Ulm, Germany
fritz.lehmann@uni-ulm.de

Daine R. Lesniak
University of Virginia
Department of Systems and Information Engineering
151 Engineer's Way
Charlottesville, VA 22904, USA
drl2v@virginia.edu

Shih-Chii Liu
University of Zürich and ETH Zürich
Institute of Neuroinformatics
Winterthurerstr. 190
CH-8057 Zürich, Switzerland
shih@ini.phys.ethz.ch

Henrik Malm
Lund University
Department of Cell and Organism Biology
Zoology Building
Helgonavägen 3
22362 Lund, Sweden
henrik.malm@cob.lu.se

Michael E. McConney
Air Force Research Laboratory
Wright-Patterson Air Force Base
Dayton, OH 45433-7702, USA
mcconney@gmail.com

DeForest Mellon Jr.
University of Virginia
Department of Biology
Gilmer Hall
McCormick Road
Charlottesville, VA 22901, USA
dm6d@ems.mail.virginia.edu

Gunnar Meyer
Universität Bonn
Institut für Zoologie
Poppelsdorfer Schloß
53115 Bonn, Germany

Richard J. D. Moore
University of Queensland
Queensland Brain Institute
St. Lucia, QLD 4072, Australia

Liviu Movileanu
Syracuse University
Department of Physics
201 Physics Building
Syracuse, NY 13244-1130, USA
lmovilea@physics.syr.edu

Peter M. Narins
University of California at Los Angeles
Departments of Physiological Science and
Ecology and Evolutionary Biology
621 Charles E Young Drive S
Los Angeles, CA 90095-1606, USA
pnarins@ucla.edu

Magnus Oskarsson
Centre for Mathematical Sciences
Lund University
magnuso@maths.lth.se

Mamiko Ozaki
Kobe University
Department of Biology
Rokkodai
Nada-ku
Kobe, Hyogo 657-8501, Japan
mamiko@port.kobe-u.ac.jp

Herbert Peremans
University of Antwerp
Active Perception Lab
Prinsstraat 13
2000 Antwerpen, Belgium
herbert.peremans@ua.ac.be

Katsiaryna Prudnikova
University of Virginia
Department of Mechanical and Aerospace
Engineering
104 Engineer's Way
Charlottesville, VA 22904, USA
kip4n@virginia.edu

Brian W. Quist
Northwestern University
Departments of Mechanical Engineering
and Biomedical Engineering
2145 Sheridan Road
Evanston, IL 60208-3107, USA
b-quist@northwestern.edu

Matthew A. Reidenbach
University of Virginia
Department of Environmental Sciences
Charlottesville, VA 22901, USA
reidenbach@virginia.edu

Filips Schillebeeckx
University of Antwerpen
Department of Mathematics and
Computer Science
Prinsstraat 13
2000 Antwerpen, Belgium
Filips.Schillebeeckx@ua.ac.be

Helmut Schmitz
University of Bonn
Institute of Zoology
Poppelsdorfer Schloß
53115 Bonn, Germany
h.schmitz@uni-bonn.de

Wolfgang Schröder
RWTH Aachen
Institute of Aerodynamics and Chair of
Fluid Mechanics
Wüllnerstr. 5a
52062 Aachen, Germany
office@aia.rwth-aachen.de

Peter Schützner
University of Ulm
Institute of Neurobiology
Albert-Einstein-Allee 11
89081 Ulm, Germany
peter.schuetzner@uni-ulm.de

Dean Soccol
University of Queensland
Queensland Brain Institute
St. Lucia, QLD 4072, Australia

Joseph H. Solomon
Northwestern University
Departments of Mechanical Engineering
and Biomedical Engineering
2145 Sheridan Road
Evanston, IL 60208-3107, USA
joe-solomon@northwestern.edu

Mandyam V. Srinivasan
University of Queensland
Queensland Brain Institute
St. Lucia, QLD 4072, Australia
m.srinivasan@uq.edu.au

Graham K. Taylor
Oxford University
Department of Zoology
Tinbergen Building
South Parks Road
Oxfors OX1 3PS, UK
graham.taylor@zoo.ox.ac.uk

Saul Thurrowgood
University of Queensland
Queensland Brain Institute
St. Lucia, QLD 4072, Australia

R. Blythe Towal
Northwestern University
Departments of Mechanical Engineering
and Biomedical Engineering
2145 Sheridan Road
Evanston, IL 60208-3107, USA
b-towal@northwestern.edu

Vladimir V. Tsukruk
Georgia Institute of Technology
School of Materials Science and Engineering
771 Ferst Drive
Atlanta, GA 30332-0245, USA
vladimir@mse.gatech.edu

Marcel Utz
University of Virginia
Center for Microsystems for the
Life Sciences
Department of Mechanical Engineering
104 Engineer's Way
Charlottesville, VA 22904
mu3q@virginia.edu

Gerhard von der Emde
University of Bonn
Institute of Zoology
Poppelsdorfer Schloß
53115 Bonn, Germany
vonderemde@uni-bonn.de

Hao Wang
University of Ulm
Institute of Neurobiology
Albert-Einstein-Allee 11
89081 Ulm, Germany

Eric Warrant
University of Lund
Department of Cell and Organism Biology
Zoology Building
Helgonavägen 3
22362 Lund, Sweden
eric.warrant@cob.lu.se

Urban B. Willi
Arlesheimerstr. 26
4053 Basel, Switzerland
uwilli@ucla.edu

About the editors

Friedrich G. Barth, born in Munich, Germany, in 1940, studied biology and human physiology at the University of Munich and at the University of California at Los Angeles. He received his Ph. D. from the University of Munich in 1967. In 1974 he succeeded Martin Lindauer as the Chair of Zoology at the University of Frankfurt am Main. In 1987 he then moved to the University of Vienna, where he established the Department for Neurobiology and founded the Austrian Neuroscience Association, for which he served as the first president. Since 2008 he has held the status of professor emeritus at the University of Vienna. His research interests centre on invertebrate neurobiology. Always considering the entire animal and its natural behaviour in its natural habitat, his approach combines fieldwork with laboratory work and the application of advanced technologies in search of functional principles. The main focus of his work has been on the workings of sensory systems and their neuroethological roles as well as related biomechanical and physical questions. Multidisciplinary collaborations with physical scientists and engineers characterize much of Friedrich Barth's research. Together with Pepe Humphrey he organized several international conferences to bring together biology and the physical sciences. The senses of spiders, remarkably sophisticated both in a biological and technical sense, and, more recently, problems of communication in meliponine ("stingless") bees took most of his attention. His own research and fieldwork, lecturing, and many guest professorships took him to numerous countries all around the world. He is a member of the Academia Europaea, the German National Academy of Sciences/Leopoldina, the Bavarian Academy of Sciences, and the Austrian Academy of Sciences. In 2001 he received the prestigious Karl Ritter von Frisch Medal of the German Zoological Society. Apart from having published some 170 full-length research papers he has served as the editor-in-chief of the Journal of Comparative Physiology A since 1996 and is the author and editor of several books. Among them are "Insects and Flowers: The Biology of a Partnership" (1991), "A Spider's World: Senses and Behavior" (2002) and "Sensors and Sensing in Biology and Engineering" (FG Barth, JAC Humphrey, T Secomb eds.) (2003).

Joseph A. C. Humphrey† was born in La Habana, Cuba, in 1948, and earned diplomas in Chemical Engineering both from the University of Barcelona, Spain in 1970 and from the University of Toronto, Canada (M. A. Sc.) in 1973. In 1977 he received a Ph. D. in Mechanical Engineering from the University of London, and in 1997 a D. Sc. in Engineering, again from the University of London. Having started out at Princeton University in 1977, Pepe (as everyone called him) joined the faculty of the Department of Mechanical Engineering at the University of California in Berkeley in 1978 and stayed there until 1994 as an Assistant Professor, Associate and ultimately Full Professor. After two years at the University of Arizona in Tucson and three

years at Bucknell University, he settled at the University of Virginia in Charlottesville in 2000 as Wade Professor of Engineering and Applied Science and Head of the Department of Mechanical and Aerospace Engineering, with an unusual joint appointment in the Department of Biology. In March 2010 his sudden death abruptly ended Joseph Humphrey's remarkable career. His main research interests were transport phenomena, laminar and turbulent flows, flow – structure interactions, and biological flows as well as bio-inspired sensors and sensing. Joseph Humphrey was a true lover of nature and enthusiastically collaborated with biologists aiming to uncover and understand flows at various interfaces and length scales and their relation to the natural environment. Among the animal sensors he mainly worked on were the medium flow sensors of spiders and fish and insect chemoreceptors. His experimental and computational research was enormously productive with some 150 archived publications. Together with Friedrich Barth, his friend of many years, he organized several international conferences bringing together engineers and biologists. Joseph Humphrey was a fellow of the American Society of Mechanical Engineers, held honorary and visiting professorships at the universities of Kyoto, Liverpool and Vienna, and served as an active member in numerous professional societies.

Mandyam V. Srinivasan has been a Professor of Visual Neuroscience at the Queensland Brain Institute and Professor of Electrical Engineering at the School of Information Technology and Electrical Engineering at the University of Queensland, Australia, since 2007. His main research interests are the principles of visual processing in simple neural systems, and the application of these principles to machine vision as used in robotics and unmanned aircraft. He has spent more than 20 years studying principles of visual flight control and navigation and the most remarkable cognitive capacities of honeybees. He has been asking questions such as "How do bees control their flight speed? How do they avoid collisions with obstacles? How do they determine how far they have flown? and "How do they orchestrate smooth landings?"

Born in 1948 and raised in India, Mandyam Srinivasan received a Master's degree in Applied Electronics and Servomechanisms from the prestigious Indian Institute of Science in Bangalore in 1970. He then moved to the United States and received his Ph. D. in Engineering and Applied Science from Yale University in 1977. Only one year later he went to Australia, where he was a research fellow of the Departments of Neurobiology and Applied Mathematics at the Australian National University in Canberra from 1978 to 1982. After a break of three years at the University of Zurich, Switzerland, he returned to Australia to work again at the Australian National University in Canberra from 1982 to 2006, since 2000 as the Director of the Center for Visual Science and since 2002 as Distinguished Professor of Visual Science. Mandyam Srinivasan is a Fellow of the Australian Academy of Science, the Royal Society of London, and the Academy of Sciences for the Developing World. Among his many awards are an honorary doctor's degree from the University of Zurich, the Rank Prize in Optoelectronics, UK (2008) and the Prime Minister's Science Prize, Australia (2006). Although formally trained as an engineer, Mandyam Srinivasan has always been interested in the interface between engineering and biology. He has published over 180 full-length research papers and edited two books to date, one of them together with S. Venkatesh entitled "From Living Eyes to Seeing Machines" (1997), and another with D. Floreano, J.-C. Zufferey and C. Ellington entitled "Flying Insects and Robots" (2009).