CHAPTER 23

Information Transfer Between Sensory and Motor Networks

M. LAPPE

Computational and Cognitive Neuroscience Lab, Department of Zoology and Neurobiology,
Ruhr-University, 44780 Bochum, Germany

Handbook of Biological Physics
Volume 4, edited by F. Moss and S. Gielen

Contents

1. Introduction

The function of the CNS is to generate and control behavior in order to support the living needs of the organism. At any time, such behavior must be appropriately adjusted to the status of the environment and of the organism itself, and hence must be guided by sensory information. The sensory and motor structures of the brain encode information in various different formats. Sensorimotor transformations interface the sensory and motor systems. They have two main tasks to solve. First, they must convert between the different coding formats. Second, they must fuse the different sensory and motor input signals and establish a unified representation of the environment and the organism's action within it.

2. Sensory and motor systems of the CNS

2.1. Sensory systems

The senses of most animals consist of vision, audition, the sense of balance, touch, smell, and taste. For the purpose of guiding motor behavior, in primates and many other animals vision, balance, and touch are most important. The associated sensations are registered by the visual, the vestibular, and the somatosensory systems. In the following, we will briefly describe these sensory systems and the primary encoding of the incoming information in each of them. In-depth information can be found in [1–4].

In vision, light entering the eye is transformed into electrical currents in the photoreceptors of the retina. After some initial image processing by the networks of the retina the visual information leaves the eye in the form of action potentials of the retinal ganglion cells. The fibers of the retinal ganglion cells form a thick parallel bundle, called the optic tract, from which branches run to several parallel retinal recipient structures. The main cortical pathway is via the thalamus to the primary visual cortex, or area V1. Other important pathways are those to the superior colliculus and to the accessory optic system and pretectum. Each retinal ganglion cell transmits information from a small localized part of the visual image, known as its receptive field. Neighboring cells transmit information from neighboring image locations. In the primary visual cortex, these topographic relationships are largely preserved (see chapter 16 by Flanagan in this book). The primary visual cortex contains a retinotopic map of the visual field. This is not true for all retinal recipient structures. While the superior colliculus is also retinotopically organized, the accessory optic system, for instance, is not. Several other order parameters are also represented in a structured manner in the primary visual cortex. These include the relative strength of input from the two eyes ('ocular dominance') and the selectivity preferences of each neuron for parameters of visual stimuli (orientation, color, motion direction) (see chapter 22 by Ernst and Pawelzik in this book). Each of these order parameters establishes a structured map in the primary visual cortex. All of these maps are present in parallel. The nature and origin of cortical feature maps are discussed in other chapters of this book.

The primary visual cortex is the starting point of information processing in a large network of more than 40 identified areas in the visual cortex [5], each containing a representation of the visual field. These areas extract and transform information from the visual image. To a first approximation, this information processing stream can be viewed as a hierarchical system for visual analysis. However, there are frequent interconnections and feedback projections between the different stages [5,6]. There is good evidence that the cortical information processing involves two major streams [4]. One is concerned with the analysis of the features of objects (shape, color, etc.) and object recognition, and maybe perception in general. The other is concerned with spatial positions of objects, visual motion, and the generation of action or motor output. Owing to the topic of this chapter, we will be mostly concerned with the latter. It involves a series of areas in parietal cortex. These areas receive not only visual input but also signals from the vestibular, somatosensory, and auditory systems and build a universal representation of space.

The vestibular organs of the inner ear sense gravitational and inertial forces and are important for our sense of balance and self-motion. Rotations and translations of the head are registered by separate sensors. The semicircular canals of the labyrinth organ are closed, liquid-filled tubes [1]. Their inner surface is lined with sensory hair cells. As the head rotates the inertial mass of the liquid generates relative movement between the liquid and the inner surfaces of the canals. This movement induces electric activity in the hair cells, which is transmitted to the vestibular nuclei in the brainstem as a signal of head rotation in space. Linear acceleration forces are sensed by the otholiths [1]. These consist of an orthogonal pair of two-dimensional arrangements of tiny calcium carbonite crystals. Inertial forces acting on these crystals are again detected by hair cells. Signals from the hair cells are transmitted to the vestibular nuclei and provide the input for the detection of linear accelerations of the head and of the tilt angle of the head with respect to gravity.

The sense of touch, the position of the limbs, the status of the muscles controlling them, the sense of temperature, and the sense of pain are all registered by the somatosensory system. For the purpose of this chapter we are mainly interested in the position of the limbs and the activity of the muscles. This is known as proprioceptive information. It is generated by spindle fibers in the muscles [1]. These fibers change their rate of action potential generation related to the length of the muscle, change of muscle length, and the contraction of the muscle. The axons of the muscle spindles of the skeleton muscles run along the spinal chord to the brain. Those from the eye muscles run along the trigeminal nerve from the face. Somatosensory information from the body is represented in somatotopic maps in the somatosensory cortex.

2.2. Motor systems

Movement of the limbs and other body parts is controlled by muscle contraction. Muscle contraction is controlled by neural commands. The neurons that control the muscles are called motor neurons. Each motor neuron is connected to a small number of muscle fibers which contract when the motor neuron fires. The motor neurons, in turn, are controlled by a network of areas from the brainstem, the cerebellum, the basal ganglia, and the cortex [1,2]. Motor neurons that activate skeletal muscles reside in the spinal chord. Motor neurons that activate eye muscles reside in the oculomotor nuclei of the brainstem.

The eye movement system or oculomotor system is the best understood of the motor systems and most of the examples in this chapter will come from the control of eye movements. It is therefore appropriate to focus here on the oculomotor system in particular. A detailed reference to the oculomotor system is [7].

The eye is moved by six muscles. One pair of muscles subserves horizontal eye movements. The remaining two pairs subserve vertical movements and rotations along the line of sight (called torsional movements). The muscles are driven by motor neurons in the oculomotor nuclei of the brainstem. Each muscle pair has its own nucleus.

Eye movements can be broadly classified into two categories, those that align gaze with a specific target and those that stabilize vision during movements of the head or body. The first category contains three important types of eye movement. The first are normal gaze shifts, called saccades, which bring the line of sight to an object of interest. These are the most ubiquitous eye movements, performed for instance several times a second while reading this sentence. The second type of gaze targetting eye movements are vergence movements. They adjust the axes of the two eyes such as to look at an object at a certain distance. Often, when gaze shifts from one object in space to another, saccades and vergence eye movements go together. The third type are smooth pursuit eye movements which are initiated when one follows a small moving target with the eyes.

The group of eye movements that stabilize vision during movements of the head or body consists of a set of three reflexes, the optokinetic, the vestibulo-ocular and the cervico-ocular reflex. They rotate the eyes opposite to the head movement such that the visual image of the world remains approximately stable on the retina. They use visual, vestibular, and somatosensory information, respectively, to infer the correct direction and speed of the eye movement.

Each type of eye movement is driven by its individual neuronal machinery. Only a limited part of that machinery is shared among some eye movements. However, all oculomotor control eventually has to go through the oculomotor nuclei in order to address the eye muscles.

3. Important concepts for sensorimotor transformations

In this section, we will visit some important conceptual issues for the understanding of sensorimotor information processing. After that, we will delve into specific topics and systems in the remaining sections. The issues discussed here shall allow to define the problems and solution strategies of sensorimotor tasks.

3.1. Coupling of action and perception

In the normal function of an organism, motor action and sensory perception are closely intertwined. First of all, perception is the basis for controlling goal-directed action. Each action, in turn, has consequences for perception because it changes the temporary relation between the organism and the perceptual world. It is possible to study motor and sensory systems in isolation in controlled laboratory conditions. Often this simplifies matters of experimental design and interpretation, much as it is easier to study simple isolated physical systems rather than complicated ones. But to truly understand either sensory perception or motor control it is necessary to consider the interaction between the two.

Sensory-guided motor action is usually an iterative procedure in which sensory input is continuously evaluated and used for control purposes while the motoric act continuously changes the setting of the actor in the environment. Think, for example, of walking through a room towards the door. Every step that takes you closer to the door also changes the view of the room that is used to locate the door and guide the movement. This iterative process of sensory perception and motor activity is called the action–perception cycle (Fig. 1).

A further aspect of the action–perception coupling is that often motor activities are performed to aid sensory perception. If you want to get a closer look at an object that is moving, smooth pursuit eye movements are initiated to keep high-resolution gaze at the object. Similarly, if you want to get information about an

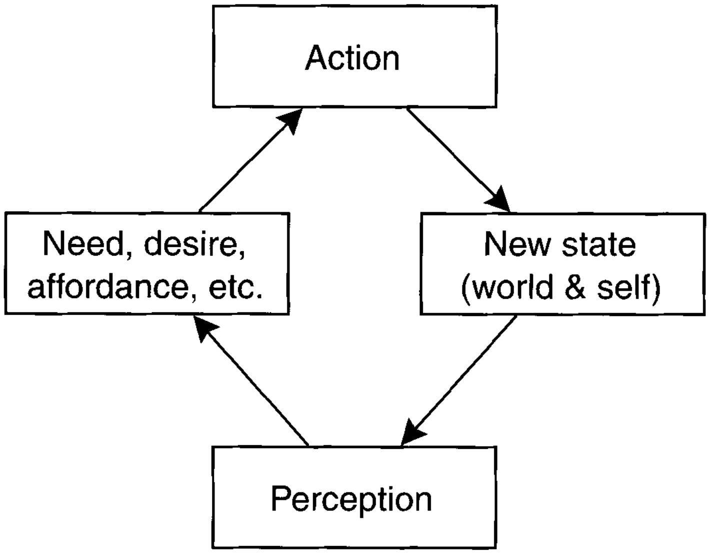

Fig. 1. The action–perception cycle.

object somewhere in the room, a saccadic gaze shift brings it to the center of the visual field. Hence it is not only true that perception provides the basis of action. Rather, action may also provide support for perception. In computer vision, recognition of this principle has led to the development of the paradigm of active vision [8].

3.2. Feedforward and feedback control systems

In motor control, the controller sends a signal to an effector that prompts the effector to perform a certain action. In biological systems, the effector is typically itself a complicated system, has its own dynamical properties, possibly exhibits nonlinear behavior, and is subject to noise and delays in internal as well as external processes. There are two principle ways by which the controller can assure that the effector performs the desired output function [9]. Either the controller knows how the effector performs, sends an appropriate command, and trusts that the effector works as expected. This is a feedforward control system. Or the controller sends a signal that approximately results in the appropriate response, checks the output, and uses the difference between the actual output and the desired output to steer the control signal towards matching the two. This is a feedback control system.

In feedforward control systems (Fig. 2A), the controller must predict the output generated by the effector. Because there is no feedback information of the actual performance, it is essential that the prediction is very accurate. There are two possibilities by which the controller can generate its prediction. Either the controller knows *what* the effector will do, i.e., directly knows what output the control signal will generate. Or the controller knows *how* the effector works and predicts the output generated by the control signal, in which case the controller is said to have an internal model of the effector [10]. The quality of a feedforward control mechanism depends on the accuracy of the prediction. In theory this could be high but in practice it is not, because biological systems behave too variably to allow an accurate prediction and normal biological behavior often has to cope with unexpected disturbances. The advantage of feedforward control, however, is that it is fast.

Feedback control systems (Fig. 2B) have less problems with variabilities of the effector and the task. This is because they do not attempt to accurately predict the behavior of the effector. Instead they use the error in the motor output to continuously adjust the control signal. The current output is fed back to the controller and compared to the desired output. The difference is used for the new control signal. This is an efficient and simple method, as it does not require much detailed knowledge about the effector or an internal model. It also has the advantage that it can deal with unexpected disturbances of the task or the behavior of the effector. However, its problems lie in the fact that: (a) the feedback signal is usually sensory in nature and hence must be interpreted or transformed into an appropriate motor error signal, and (b) the error signal arrives with a certain temporal delay, as it needs to be registered and processed by the sensory system first. Because of this, a

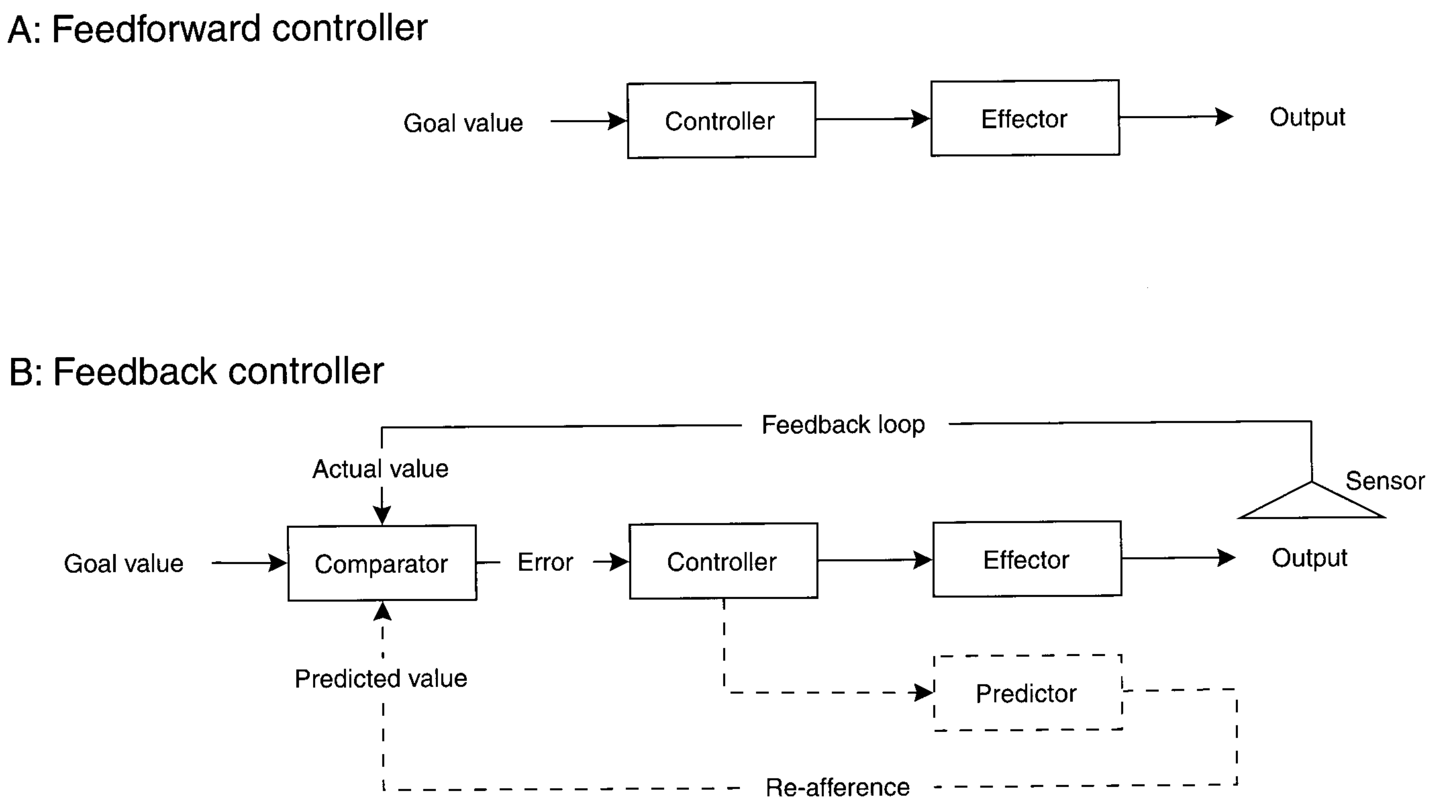

Fig. 2. Feedforward and feedback control systems.

feedback controller typically lags behind its goal and only asymptotically reaches perfect performance. Moreover, if the feedback controller is not able to follow correctly it develops a phase lag and might become unstable.

A means to cope with these problems is to add a faster way to generate an error signal. This can be done by using an internal model of the effector to predict the error (Fig. 2B). This internal model receives the same control signal as the effector and without delay generates an expected output signal. This expected output is fed back to the controller and is used like the true error signal that later arrives via the regular sensory feedback mechanisms. In biological systems this is called the re-afference method and the signal that is sent to the model of the effector is called the efference copy signal [11,12].

A feedback control system consists of a closed loop of control and error signals. If the feedback path is cut the system is in the so-called open-loop situation. In this case, the system behavior is entirely determined by its feedforward path. This makes it possible to study the behavior of that part in isolation. An often used technique, for instance, is to look at only the initial phase of the systems reaction to an input. Because of the temporal delay in the sensory feedback the initial response phase is of the open loop type. The performance of the system is specified by its gain, which is defined as the ratio of the output of the system to the input. In the closed loop situation the gain should ideally be unity, which is often approximately reached. Yet the gain normally depends on the frequency distribution of the input signal. The open loop gain is a measure of the behavior of the feedforward path of the system only. It is often very different from the ideal value of 1.0, which is desirable as it affects the speed of adjustments of the control signal by the feedback path.

3.3. Plasticity of input–output relations

The appropriate motor reaction to a sensory input may vary greatly in everyday circumstances. Moreover, the exact motor program that leads to a desired goal varies depending on constraints experienced in the situation. For instance, obstacles in the path may force deviations from a straight trajectory. Changes to the sensory input (think for instance of wearing astigmatic glasses) or to the effector (for instance by an additional load to be moved) require a recalibration of the mappings from sensory to motor coordinates. For these reasons sensorimotor transformations cannot be fixed but have to allow plastic changes to adapt to changes in the behavioral situation or the sensory input.

Most sensorimotor behaviors show the capability to adapt to changes. In fact, many sensorimotor transformations are under constant control by recalibration procedures. This recalibration is performed through a continuous comparison of the desired motor effect with the actual motor output. Depending on this comparison, the connections and weights of the sensorimotor transformation are adjusted. Very likely feedback pathways from cortical and subcortical control centers through the cerebellum are involved in the constant recalibration process [10,13].

3.4. Multiple frames of reference and sensor fusion

Sensory information is encoded in different primary formats in the different sensory systems. For example, in the visual system the position of an object is encoded spatiotopically in the distribution of activity in the two-dimensional layer of retinal ganglion cells. In the auditory system the spatial position of that same object is encoded in time and intensity differences between the two ears. On the output side, different motor actions that involve different effectors might also be coded differently. Saccadic eye movements are encoded by the spatial distribution of activity in a two-dimensional topographic map in the intermediate layers of the superior colliculus. The encoding of arm movements in the primary motor cortex, in contrast does not appear to have a topographic organization.

To transfer sensory information to appropriate motor commands hence requires to interconvert between different encoding schemes. Because the same information might be used for many different purposes it is required to keep it in several different formats in parallel. For instance, the visual location of an object might be used to direct an eye movement to it, or to reach at it with the arm. Encoding information such that it can be used in multiple output formats is achieved in two ways. On the one hand, the brain constructs multiple separate encodings each serving a different purpose and residing in a different brain area. On the other hand, the brain uses encoding schemes that are more universal and that provide the encoded information in such a way that a brain area that requires a specific format may extract it by specific read-out mechanisms of the universal code. This will be elaborated in the next section.

A further complication is that the different primary sensory and motor encoding formats do not have a fixed relationship with one another. Neither do they have a fixed relationship to the external world. This is because of the coupling of action and perception. Every action changes the orientation of the sensors in the world and with respect to each other. For instance, eye movements change the orientation of the eye in the head such that the same location on the retina now corresponds to a different direction relative to the head. Head movements likewise introduce a dissociation between spatial locations relative to the head and the trunk. This also applies across sensory modalities. Since the ears are fixed in the head, eye movements also introduce a dissociation between the encoding of the visual and of the auditory location of an object.

It is therefore appropriate to ask in which frame of reference, i.e., with regard to which coordinate system, information is encoded. We can distinguish between five frames of reference. Retino-centric encoding is with respect to the retina or the eye. Head-centric encoding is with respect to the head. Body-centric is with respect to the trunk or body. Ego-centric encoding gives the location in external space with respect to the current location of one self. Allo-centric encoding means the location of an object in the world, irrespective of the position and orientation of the body or the sensors. This is what we commonly perceive as the locations of the objects around us. Sensorimotor transformations must keep the encodings of information in these various frames of reference in continuous register.

Although primary sensory information comes in different formats it often pertains to the same object or the same quality. Spatial position can be sensed visually, acoustically, or haptically (i.e., by means of touch). Movement can be sensed visually, acoustically, haptically, or vestibularly. An object's identity is conveyed by its shape, color, feel, smell, or taste. To use all sensory information as efficiently as possible requires to combine or *fuse* them, in order to form a coherent percept or generate a successful motor command. Such sensor fusion, in turn, requires a universal encoding scheme or representation that is beyond the primary sensory encodings in the different formats. This unified, supramodal encoding of space is established along the pathways from sensory to motor areas of the cerebral cortex.

3.5. *Distributed encoding in overdetermined, noncartesian coordinate systems*

Mathematically, spatial transformations can most simply be described by vector addition, rotation, and scaling in an orthonormal coordinate system that has the smallest possible number of degrees-of-freedom. In the brain, things are different. Spatial parameters are often encoded in the firing rates of large populations of neurons. Each neuron contributes only a small part to the entire encoding of a single parameter and might participate in the encoding of several other parameters as well. A well-known example is the population encoding of movement direction, for instance the direction of arm movement in primary motor cortex or the direction of visual motion in motion sensitive visual areas. In the population code, the direction of movement **D** is represented by a vector-weighted summation of the activity of all neurons in the population

$$\mathbf{D} = \sum_{i=1}^{N} a_i \mathbf{e}_i, \tag{1}$$

where a_i is the firing rate of neuron i, and $\mathbf{e}_i$ is a unit vector in the direction in which the neuron is assumed to contribute to the encoding. Usually this is the direction for which the neuron has the strongest firing rate, i.e., the preferred direction of the neuron. Since every neuron has its own preferred direction, and since the preferred directions are often equally distributed across the neuronal population, Eq. (1) is a linear decomposition of **D** in a vastly overcomplete set of basis vectors.

Under the constraints imposed by properties of a biological system, such a distributed population encoding has a couple of advantages. Among them are the robustness against noise or against failure of individual neurons, and the ability to perform smooth interpolations. More information on this is provided in chapter 20 by Gielen and chapter 19 by Treves. With regard to the topic of this chapter, there is one additional important property, namely that it is possible to represent several different encodings with the same population of neurons. Consider, for instance, a population of neurons encoding visual motion in a specific part of the retina. Each neuron has its own preferred direction $\mathbf{e}_i$ and speed s_i of motion. Then by using the above population code

$$\mathbf{D} = \sum_{i=1}^{N} a_i \mathbf{e}_i,$$

one might retrieve the direction of visual motion, as the speed of visual motion is averaged out by summing over all neurons. This assumes, of course, that the combination of speed and direction tuning is equally distributed across the population. On the other hand, one might construct a similar population code for the speed S of the visual motion by weighting each neuron by its preferred speed

$$S = \sum_{i=1}^{N} a_i s_i.$$

Or one might retrieve the full velocity $\mathbf{V}$ by using both the direction and the speed in the population code

$$\mathbf{V} = \sum_{i=1}^{N} a_i s_i \mathbf{e}_i.$$

This argument may seem trivial, but the point is that the neuronal population provides the complete information and the process by which the population is read out determines what information is used and how it is used. In the brain, this means that subsequent areas can each choose to select different parts (or different formats) of the information that is provided by the preceding area. This is important for the construction and use of a supramodal representation of space as outlined in the previous section.

The argument becomes more interesting if we consider it in relation to such a supramodal space representation and to the issue of different frames of reference. Let us assume that the population of neurons described above also receives information about the orientation of the eyes in the head, either from the stretch receptors in the eye muscles or via an efference copy signal. Then in an analogous way it is possible to retrieve by different population read-out procedures the visual motion on the retina (a retino-centric variable), the position of the eye in the head, or the visual motion with respect to the head (a head-centric variable). Now this is exactly what would be required of a mechanism for sensorimotor coordinate transformation. In fact, this is one of the ways by which sensorimotor transformation is realized in the brain.

3.6. Separation of state variables position and velocity

Physically, position and velocity are closely related as velocity is the derivative of position. The brain often treats them as two separate and unrelated entities. For instance, it is possible to perceptually experience motion without a change in position. This is demonstrated, for example, in the motion aftereffect [14]. If one looks at a continuous motion pattern for about a minute and then suddenly looks at a stationary scene the scene appears to move while at the same time it does not appear to change position. Another example of the separation of position and velocity can be seen in the control of smooth pursuit eye movements. When a fixated stationary object suddenly starts to move the brain initiates two superimposed eye movements to follow the motion of the object. First there is a smooth acceleration of the eye such as to bring its speed up to the speed of the target. While this acceleration occurs, a saccadic gaze shift is initiated to bring the target back into the center of gaze. These two movements occur independently and at different times after the onset of the target movement, and are controlled by different systems.

The separation of position and velocity is also seen in the functional anatomy of the brain. There is a dedicated processing stream that explicitly involves motion. This processing stream analyses visual motion and provides the motor commands for motion-related motor acts such as tracking eye movements or the control of locomotion and posture. The analysis of spatial position and the control of motor acts directed towards spatial positions, i.e., saccadic gaze shifts and reaching and pointing movements, is largely subsevered by a different network of brain areas. In the following, sensorimotor transformations for spatial actions are considered first and then the processing of motion.

4. Spatial representations and transformations

4.1. Topographic representation in early visual areas

In an abstract sense, the eye functions similarly to a camera, in which the optics (lens) generate a faithful two-dimensional image of the visual scene on the receptor surface (the retina). Adjacent photoreceptors receive light from adjacent visual directions. The topography of the visual image is preserved in the different layers of the retina and in several of the retinal recipient structures in the brain, such as the superior colliculus and the lateral geniculate nucleus. It is also preserved in the primary visual cortex, or area V1, and subsequent visual cortical areas V2, V3, V4, V5 (or area MT), and V6 in the cortical processing stream. In each of these structures, a neuron can be characterized by its retinotopic receptive field, i.e., the part of the retina from which this neuron receives information. Receptive fields become bigger along the path, increasing in diameter from less than a minute of arc in the retina up to 10° of visual angle in the peripheral visual field representation in area MT. In the central part of the visual field receptive fields are smaller than in the periphery and the number of neurons per degree of visual angle is higher. Thus more cortical tissue is devoted to the processing of the central parts of the image. This is known as the cortical magnification factor. However, while this affects the metric of the representation, there is a strong tendency in the early visual areas to preserve the topographical relationships between image positions such that neighboring neurons have adjacent or overlapping receptive fields. More information on cortical maps is provided in chapter 22 by Ernst and Pawelzik.

In the most simple form, the receptive field of a visual neuron can be modeled as a Gaussian with its center in the receptive field center and its width adjusted to the width of the receptive field (Fig. 3A). For many models that deal with spatial

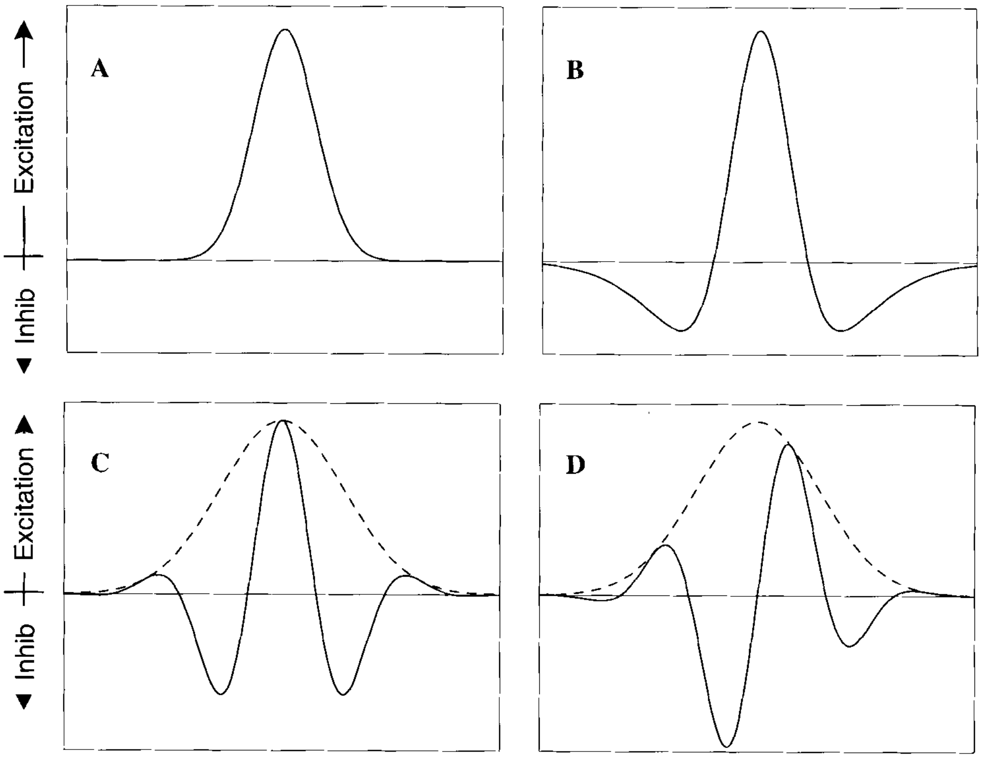

Fig. 3. Descriptions of receptive fields. (A) Simple Gaussian profile. (B) 'On'-center, 'Off'-surround structure obtained as the difference of two Gaussians. (C) and (D) Gabor functions with phase 0 (C) and phase π (D).

position and the conversion between coordinate frames this is already sufficient. However, the receptive fields of most visual neurons show a richer structure which is related to their respective role in the processing of visual information. Already in retinal ganglion cells the receptive field consists of two parts. One part leads to excitation of the cell when it is hit by light, the other part leads to an inhibition when illuminated. These are known as the 'on' and 'off' regions of the receptive field. In the retinal ganglion cells, and also in the neurons of the lateral geniculate nucleus, the on and off regions are arranged as two concentric circles of different diameters. They can be described by two Gaussians of different widths and signs. The total receptive field is then described by the difference of the two Gaussians (Fig. 3B) [15,16]. In the primary visual cortex and beyond, receptive fields become more complex in correspondence to the more complex response properties of the neurons. Neurons in V1 respond selectively to the orientation and spatial frequency of gratings presented in their receptive field. The receptive fields of such neurons can be modeled by Gabor functions, i.e., the product of a Gaussian and a cosine function (Fig. 3C,D) [17,18].

4.2. Construction of three-dimensional space

The retinal image of each eye is two-dimensional. Accurate goal-directed spatial action requires knowledge of target position in three-dimensional space. The brain, therefore, must construct a representation of the three-dimensional world. There are many cues to the third dimension already in the two-dimensional monocular image (Fig. 4). Among these cues are shading, textural density and perspective, object-size relationships, and motion parallax [3,11,19]. Motion parallax is the differential visual motion that objects in different distances from the eye undergo when one moves the head sideways, for example (Fig. 4C). All of these cues are evaluated by the visual system and are used in parallel to reconstruct three-dimensional spatial relationships. Their usefulness for depth perception varies with the depth scale that an individual cue may provide and with the depth range over which it can be analyzed [20].

A primary cue to depth in the near range (below 10 m) originates from binocular vision [21]. Each eye sees the world from a slightly different perspective. From the difference in perspective, parallactic differences between the images of the two eyes result (Fig. 5). They are called horizontal retinal disparities. In Fig. 5, the lines of sight of the two eyes are converged on the point F and form a certain *vergence angle* γ_F. In both eyes, point F is projected onto the center of gaze. In contrast, for the

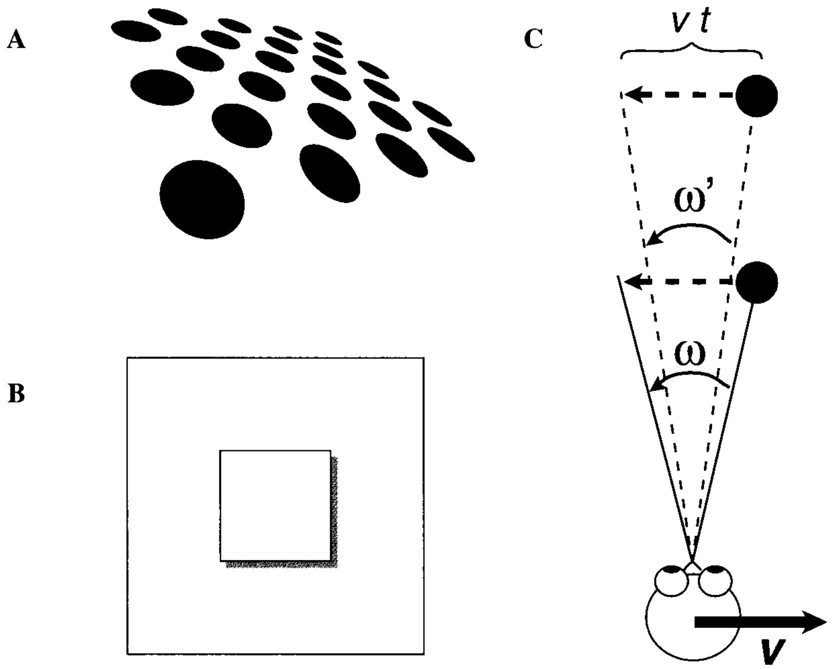

Fig. 4. Monocular cues to three-dimensional structure. (A) Size foreshortening, texture density, texture perspective. (B) Shading. (C) Motion parallax.

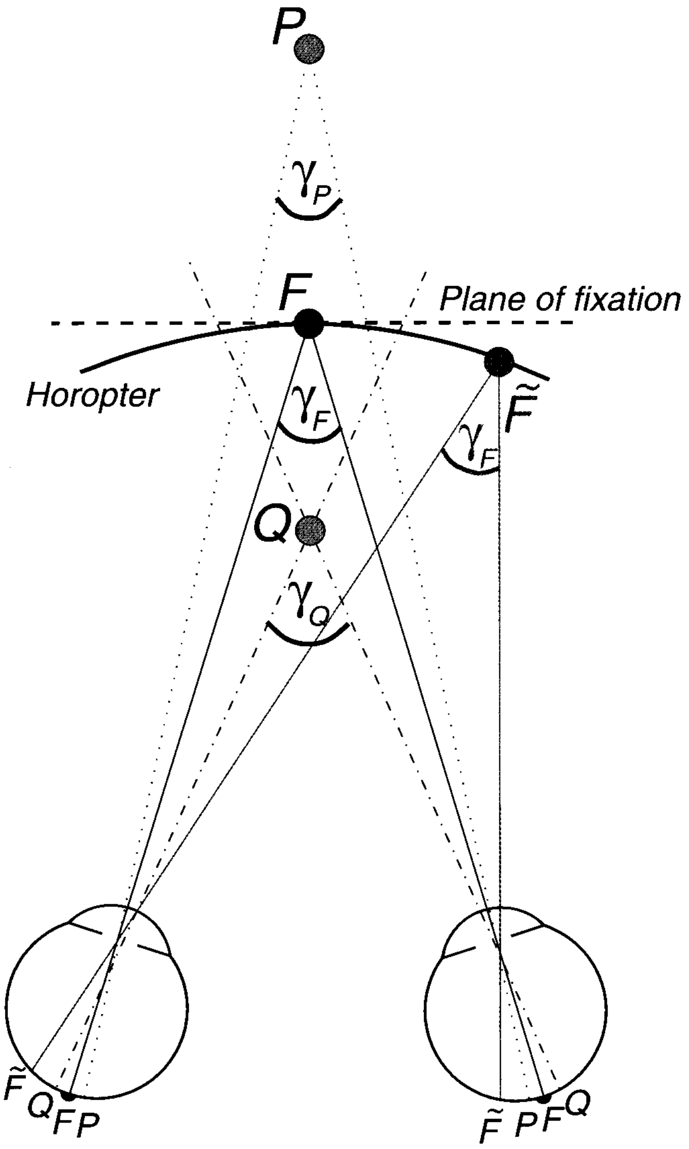

Fig. 5. Binocular horizontal disparity.

more distant point P, the rays through the center of the lens of the eye form an angle γ_P. The difference between the two angles,

$$\delta_P = \gamma_P - \gamma_F$$

is the absolute horizontal disparity of P. Its value can be used to estimate the distance of P from F. The absolute horizontal disparity is negative when P is more distant than F. The absolute horizontal disparity of point Q, $\delta_Q = \gamma_Q - \gamma_F$, in contrast, is positive. This point is closer to the eyes than F. A special set of points $\tilde{F}$ consists of those for which the absolute disparity is zero. They form a two-dimensional curve that is called the horopter.

The disparity of a point can be obtained from the projection of the image point in the two eyes. The projection of F is in the center of gaze of both eyes. The projections of a point $\tilde{F}$ on the horopter fall on eccentric positions in both eyes. The two positions of the projection of $\tilde{F}$ are called 'corresponding positions' because they both correspond to the same object with zero disparity. In contrast, the projections of point P do not fall on corresponding positions. In the left eye, the projection is to the right of F. In the right eye, it is to the left of F. The angular difference between the two projections of P is equivalent to the absolute horizontal disparity δ_P of P.

Once the absolute horizontal disparity δ_P is known, recovery of the true distance of point P from the eye further requires knowledge of the distance between the eyes, and of the vergence angle γ_F. Only if all of these parameters are known, the distance of P from the eye can be calculated geometrically. However, from the visual images in the two eyes solely the disparity can be retrieved. The vergence angle and the interocular distance cannot be determined visually. Hence, an absolute depth judgement is not possible from the visual information alone. Moreover, for geometric reasons, the calculation of absolute depth strongly relies on the accuracy of the vergence angle measurement. Any errors in that measurement will lead to large errors in depth perception. However, the difference between the disparities of two points P and Q,

$$\delta_{PQ} = \delta_P - \delta_Q = \gamma_P - \gamma_Q$$

can be retrieved from the binocular images alone. This difference is called the relative horizontal disparity between P and Q. The relative horizontal disparity is independent from the vergence angle. It permits a direct visual estimation of the depth difference between two objects solely from image information.

In the primary visual cortex and in several higher cortical areas, the majority of neurons receive visual input from both eyes. Many of these neurons are selective for disparity [22]. Their response to a visual stimulus depends on the binocular disparity of that stimulus. These neurons are considered to form the basis of our stereoscopic depth perception. Models of disparity sensitive receptive fields have proposed two different mechanisms [23,24]. Either the neuron receives input from two different (noncorresponding) retinotopic locations in the two eyes [25]. Models of this type are called position-based models. Or the neuron receives input from two corresponding locations, but with a different phase of its Gabor function in the two eyes [26–29]. Models of this type are called phase-based models. In phase-based models, the neuron becomes sensitive to differences in the horizontal position of a textured object, for instance, because the left eye might be optimally stimulated when the image of the texture is at phase zero, while at the same time the right eye is optimally stimulated by a texture that is shifted a bit, i.e, at a phase different from zero. It is also possible to combine both approaches and arrive at a hybrid model [23]. However, while the phase-based and the position-based models start from different assumptions, have different physiological relevances, and involve different computational steps, it can be demonstrated that at the final stages where disparity values are made explicit, the simplest versions of the two methods are mathematically equivalent [30].

From the responses of binocular disparity-sensitive neurons it is possible to infer the relative depth between two visible points. Lehky and Sejnowski [31] have shown how a population code with physiologically plausible parameters can account for human disparity sensitivity data and depth interpolation. To estimate true egocentric distance, absolute disparities and a signal describing the vergence angle are necessary. Such a signal is provided by a modulation of the firing rate of individual binocular neurons in area V1 [32] and in the parietal cortex [33]. These neurons are not only selective for disparity, they are also influenced by the vergence angle of the eyes. This combined selectivity can be used to establish a distributed representation of egocentric distance [34].

4.3. Multiple space representations in parietal cortex

Retinotopic receptive fields, even binocular ones, define the position of an object in space in a retinal frame of reference. Spatial position is encoded relative to the current direction of gaze. Many spatial actions, however, require an encoding of target position with respect to the body or the external world. To direct an arm movement to the correct point in space, for instance, the brain needs to have information about the position of that object relative to the shoulder.

Knowing the location of an object in retinotopic coordinates is not enough to specify its location in body-centric coordinates because the eyes can move relative to the body. Thus, retinal position information needs to be combined with information about the position of the eyes in the head to create a representation of the object's position in head-centric space. If this is then combined with information of the position of the head on the body a body-centric representation becomes possible. Most research up to now has focussed on the first transformation, that of retinocentric representations to head-centric representations. The second step, head-centric to body-centric has only lately been explored in some detail.

Current concepts of space representation in primate cortex center around three neurobiological findings. The first are spatial gain fields that modulate the response to a visual stimulus depending on eye position in the head. Second, neurons with head-centric receptive fields have been found in some brain areas. Third, some neurons shift their receptive field dynamically to a future retinotopic location before an impending eye movement. These mechanisms may act in parallel, or might depend on each other in a serial or circular fashion. Several observations indicate that they could form multiple parallel mechanisms rather than a single hierarchical process. We will look at each of them in turn.

4.3.1. Implicit distributed coding by spatial gain fields

Neurons in many areas of the brain scale their response to a visual stimulus within their receptive field depending on the current position of the eye in the head [35].

They respond to a preferred stimulus only when its image falls on a specific retinal location. However, the response strength changes when the eyes (and the stimulus) move to a different direction in space. Thus, while these neurons have a clearly retinotopic receptive field they also carry information about the current eye position. The influence of eye position on the activity of neurons has been termed the 'spatial gain field'. It has been described for many areas along the processing stream towards and within parietal cortex (areas V3A, V6A, MT, MST, LIP, and 7A) [36–39] but also for premotor cortex [40] and superior colliculus [41]. The widespread occurrence of eye position gain fields in the monkey brain could suggest that they subserve a basic form of space coding. The origin of the modulatory input is unclear. It could be proprioceptive feedback from the eye muscles, or a copy of the motor command to move the eye, or a combination of both.

Several theoretical studies suggested that gain fields may serve to transform the coordinates of the incoming sensory signals to a non retino-centric representation of space. Zipser and Andersen [42] developed a backpropagation network that used an extraretinal eye-position signal to transform retinotopic visual input into a head-centric representation. The network consisted of three layers of neurons. The input layer contained neurons with retinotopic receptive fields modeled as Gaussians. A second set of inputs encoded the position of the eye in the head. These neurons increased their firing rate linearly with eye position in a preferred (horizontal or vertical) direction. The output layer was set to encode head-centric position of a target. Training data consisted of combinations of retinal input, eye position, and the corresponding head-centric output. The network was trained to associate the correct input and output patterns with a backpropagation learning rule. The units in the intermediate 'hidden layer' developed retino-centric receptive fields but their activity was also modulated by eye position. Their behavior was functionally similar to the gain field neurons in area 7A, suggesting that the role of these neurons might lie in a transformation between reference frames. Later studies have refined the general ideas of Zipser and Andersen by using more biologically plausible learning mechanisms [43] and examining the consequences and function of head-centered coordinates in more detail [44,45]. Bremmer et al. [46] showed with real experimental data that a population of neurons is capable of a coordinate transformation of visual signals into a non retinocentric frame of reference.

Pouget and Sejnowski [47] have formalized the spatial transformations provided by gain fields in the theory of basis functions (Fig. 6A). This formalism capitalizes on the fact that any smooth function can be expressed in a series of basis functions. Classical examples are the Taylor series or the Fourier series. Likewise, a set of Gaussians with different centers and widths, or a set of sigmoids with different centers and slopes, also form a basis set to express any smooth function in a series [48,49]. Pouget and Sejnowski described the receptive field of a single neuron by a Gaussian and the eye position gain field by a sigmoid. Both interact multiplicatively such that the behavior of the neuron is described by the product of a Gaussian and a sigmoid (cf. [50]). This product, in turn, also defines a basis set, provided that all possible combinations of parameters are included [47]. Hence each neuron can be interpreted as providing the amplitude of a single basis function from that set.

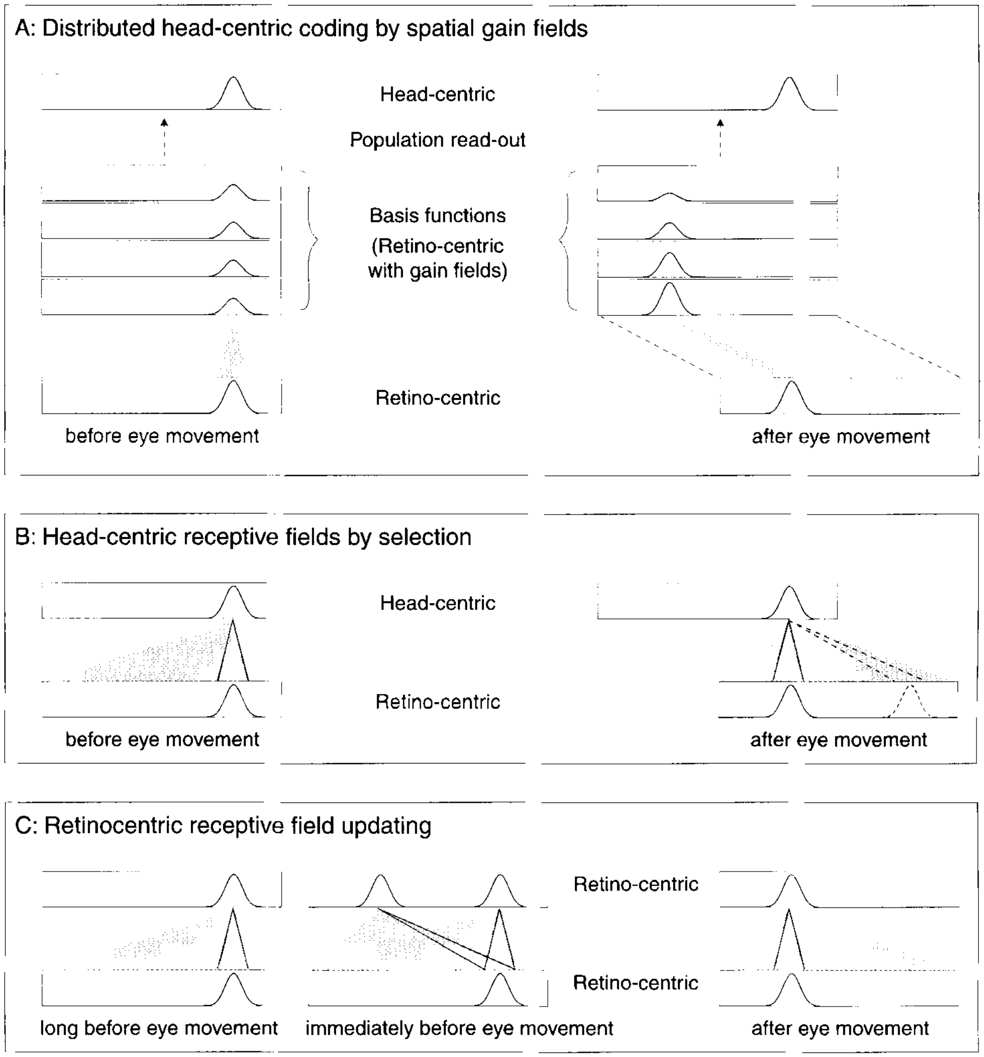

Fig. 6. Schematic illustrations of the three types of dynamic spatial representation described in Sections 4.3.1 and 3. Gaussian functions depict areas of excitation in representational maps. Shaded areas represent range of neural connections. See text for details of the mechanisms.

Equivalently, each neuron can be interpreted as encoding the input at a specific retinotopic location (its Gaussian receptive field) and a specific eye position (its sigmoid gain field). By weighted summation of all neuronal activities with appropriately chosen weights it is possible to represent any smooth mapping from input (spatial position and eye position) to a desired output. The weights, in that terminology, correspond to the amplitudes of the different basis functions in the expansion series. Specifically, it is possible to choose weights such that the output becomes a Gaussian function of head-centric position. In this case, the neurons encode head-centric position. However, they do so in a distributed, i.e., implicit manner. The appropriate weights can be found by learning procedures. Pouget and Sejnowski [47] used the delta-rule to learn the mapping from retinotopic to head-centric encoding. The delta rule minimizes the squared error between the actual output and the desired output by gradient descent. Also simpler, correlation-based learning methods are feasible [51].

The principle of implicit distributed encoding is schematically illustrated in Fig. 6A. In the left and right drawings of Fig. 6A the same head-centric target is seen from two different eye positions. In the left drawing, the eye is in the central position and the target falls on the right retinal hemifield, where it elicits a Gaussian activity profile. This activity is transmitted to the basis function encoding in the parietal cortex, where it excites all neurons that have connections with this part of the retina. Because the eye is in the central position no gain field modulation occurs and all responses have the same amplitude. In the right drawing of Fig. 6A, the eye is shifted to the right. The image of the target now falls on the left retinal hemifield. This excites a different subpopulation of parietal neurons, namely those that have receptive fields in the left retinal hemifield. However, the distribution of activity within the subpopulation is nonuniform. Because eye position is eccentric, the gain field modulation leads to strong responses in neurons in which the gradient of the gain field is in eye movement direction. Examples are the lower two of the four basis function neurons in Fig. 6A. In contrast, neurons for which the gradient of the gain field is against the eye movement direction exhibit only weak responses (upper neuron in Fig. 6A). Hence the amount of excitation varies within the population of neurons because each neuron is modulated differently by eye position. The distribution of activity in the subpopulation defines how the retinotopic position of the target corresponds to head-centric coordinates. The population read-out mechanism interprets the activity distribution in the population and establishes the location of the target in head-centric space. For an equal distribution of activity in the subpopulation of excited neurons, the head-centric location is identical to the retinocentric position (Fig. 6A, left). For an asymmetric distribution of activity, the head-centric position of the target must be shifted depending on the degree of asymmetry (Fig. 6A, right).

An advantage of the implicit, distributed representation by basis functions is that it can be used to encode not only one particular transformation, but any other input–output mapping as well. The particular transformation depends on the weights that are used to read out the population activity. In that sense, this type of encoding is coordinate-free (the coordinates are chosen when the weights are defined) and can represent or generate different spatial representations in parallel. Salinas and Abbott [51] have demonstrated that it is possible to interface such a representation directly to the population coding of motor output in primary motor cortex (see Section 5.3). This requires a different set of weights than the extraction of head-centric position, but it can be subserved by the same population of gain-field neurons. Van Opstal and Hepp [41] have shown how the parameters for the control of goal-directed saccadic eye movements (Section 5.1) can be obtained from such a representation.

4.3.2. *Explicit coding by head-centric receptive fields*

The above model assumes that head-centric position of objects is encoded implicitly in a population of neurons. While this type of encoding has certain advantages in terms of flexibility it may, on the other hand, sometimes be unwieldy to use because the responses of an entire population of neurons must be combined before a true head-centric signal can be reached. A more direct way to represent head-centric position would be to explicitly construct head-centric receptive fields. Indeed in a small number of cortical areas neurons are found that possess head-centric receptive fields. These receptive fields remain in head coordinates even when eye position changes. Head-centric receptive fields have been observed in area V6A [37], area VIP [52], and premotor cortex [53]. However, the existence of cells coding explicitly in a head-centric frame of reference does not exclude eye position effects in the same area or even in the very same cells. About half of the cells in areas VIP, V6A and the premotor cortex reveal eye position gain fields on their firing rate [37,52,54], both in darkness and in normal viewing. Interestingly, in area VIP, this eye position effect occurred in both eye-centric cells and head-centric cells.

Gain field models have shown that head-centric receptive fields can be constructed from a combination of retino-centric receptive fields and spatial gain fields in a hierarchical fashion [42,45]. But they have also suggested that the explicit head-centric step is not necessary, since all information is present implicitly in the neuronal population [42,47,51]. A strict hierarchical construction of head-centric neurons from spatial gain fields would predict that once head-centric neurons are established, gain fields are no longer needed. Yet gain fields are found in many cortical areas in the dorsal stream. Moreover, even the head-centric neurons in area VIP themselves (the putative end point of a hierarchical construction) exhibit eye position gain fields [52]. This might reflect a residual effect of the construction by gain fields. But it could also mean that head-centric receptive fields are generated by a different mechanism. For instance, dynamic selection of input from a retinotopic representation could directly yield head-centric receptive fields (Fig. 6B). In this view, a head-centric neuron makes connections to the entire representation of the retina (shaded areas in Fig. 6B) but selectively gates its connections so as to restrict its input to only a part of the visual field (continuous vs. stippled lines in Fig. 6B). The selection, then, is adjusted based on eye position. In this view, the observation that head-centric neurons in VIP are modulated by gain fields would suggest that gain fields have a functional importance that goes beyond the construction of head-centric receptive fields. In fact, gain fields can also be observed in area LIP [39,55], where a third mechanism for spatial localization, receptive field updating, has been described (see below), indicative of a parallelism between the different spatial coding mechanisms.

4.3.3. *Dynamic retino-centric receptive field updating*

Neurons in some brain areas have retinotopic receptive fields that anticipate the effect of impending eye movements. Slightly prior to a saccadic gaze shift, they shift their receptive field in space to the position that will be retinotopically correct after the saccade is completed. These phenomena were first observed in LIP by Duhamel et al. [56], and later in the frontal eye field [57], and superior colliculus [58]. While these neurons appear to encode a certain retinotopic location, their receptive field cannot be simply anchored to input from only that retinal location. Rather, they must receive information from a much larger area of the retina and dynamically evaluate only a restricted part of the input. A possible mechanism may be a spatially and temporally variable, gaze-dependent gain modulation of the receptive field structure (Fig. 6C). This is similar to the model of direct construction of head-centric receptive fields outlined in the previous section. However, unlike in head-centric neurons, which continue to use the information from the new area of the retina when the eye is in the new position, the retino-centric neurons only transiently use that area but switch back to the original part of the retina that corresponds to their retinotopic receptive field immediately afterwards (Fig. 6C). In the intermediate step, the neuron can be driven by stimulation of either the old or the new retinal location.

For anticipatory receptive field shifts, information about the direction and amplitude of an impending saccade are needed. Quaia et al. [59] have proposed a model of receptive field updating. This model takes into account the latencies and firing properties of the neurons in area LIP, the frontal eye field, and the superior colliculus. It assumes that information about impending eye movements is provided by the oculomotor signal from the frontal eye field that precedes the eye movement.

4.3.4. Why multiple space codes?

Why does the brain apparently use multiple parallel space codes? Certainly the different encoding schemes differ in a number of behaviorally relevant properties such as accuracy, flexibility, robustness, number of neurons required, demands on the structure of the input or output, etc. This means that each mechanism has a certain functional scope for which it is optimal, or at least for which it is superior to the other mechanisms. For instance, it would seem that an explicit head-centric receptive field must in principle obtain information from every place on the retina as the eye position changes, requiring a heavy convergence of inputs on every head-centric neuron (Fig. 6B, shaded areas). This is costly in terms of connectivity. In contrast, retinotopic neurons with spatial gain fields need only connect to those parts of the retina that are within their receptive field (Fig. 6A, shaded areas). This is more efficient than head-centric receptive fields, particularly in early visual areas where retinotopic receptive fields are small. On the other hand, even though head-centric receptive fields are more expensive, they might be preferable when it comes to accuracy or robustness, or when a specific output format is required. Also, they might provide head-centric information faster because the population summation step is omitted.

It is also conceivable that certain functions or tasks require particular encoding strategies in appropriate reference frames. Theoretically all relevant information for spatial coding can be provided by any of the three coding mechanisms, and easily transformed from one to another. Yet in many areas there seems to be more than one mechanism at work. Areas VIP and V6A have gain fields and explicit head-centric receptive fields in the same neuron [37,52]. LIP neurons have explicit retinotopic receptive fields, which they update prior to a saccade [56], but they also show eye position gain fields [39,55]. This shows that multiple space encodings are available in parallel, and that whichever encoding best serves the functional task of an area becomes the explicit code. Explicite head-centric coordinates might be preferable for the control of reaching movements of the arm (area V6A), for instance, or for a multimodal representation of body surfaces (area VIP). Retinocentric representations, in contrast, could be preferable for the control of eye movements (area LIP).

5. Goal-directed spatial action

5.1. Saccadic gaze shifts

Saccades are rapid eye movements that align the direction of gaze with a particular target of interest. Saccades are fast (up to 600°/s), quick (lasting about 50 ms), and frequent (we perform about three of them every second, mostly without ever noticing). Saccades are probably the most ubiquitous but also the most simple form of goal-directed behavior. For this reason, they have been studied extensively and have served as a prime example of the basics of sensorimotor information transfer.

Natural saccadic gaze shifts are usually a combination of an eye-in-head movement and a head-on trunk movement. However, most work on saccades has focussed on a situation where the head is fixed and only the eyes move. Only recently have researchers begun to investigate natural eye–head gaze shifts. Most of what follows in this section pertains to the head-fixed situation.

Like all eye movements, saccades are ultimately generated by contraction of the eye muscles. Contraction of the eye muscles is governed by the firing of neurons of the oculomotor nuclei in the brainstem. This final output part of eye motor control is shared by all classes of eye movements. The control systems that act before the oculomotor nuclei are specific for each type of eye movement. While saccades may appear to be rather simple movements, their control involves a large network of subcortical and cortical areas. To understand this network and the flow of information within it, it is useful to consider three separate stages: saccade planning and preparation, saccade initiation, and saccade execution.

Saccade planning and preparation refers to the process of choosing a target for a saccade, calculating its position in space, and relating that information to the saccade initiation system. Area LIP in the parietal cortex provides information about salient visual objects that can become targets for a saccade [60]. The visual receptive field of many LIP neurons serves a dual function as a motor field for saccades. Electrical stimulation in a cluster of LIP neurons with similar receptive field position initiates a saccade in the direction of that position [61]. Hence the visual retinotopic map in LIP can also function as a spatial motor map for saccades. A few other cortical areas are also involved in saccade planning and preparation, most importantly the frontal eye field (FEF) [62]. Similarly to LIP, electrical stimulation in FEF also generates saccades with a particular amplitude and direction.

Information about the spatial location of a saccade target is relayed from LIP and FEF to the superior colliculus (SC), which is a primary structure for saccade initiation. The SC contains a retinotopic motor map for saccade generation (Fig. 7A). Each location in this map is associated with a particular direction and amplitude of a saccade. Electrical stimulation at a specific map position leads to an eye movement with the respective direction and amplitude (e.g. points a and b in Fig. 7A). The spatial parameters of a saccade are retrieved from the implicit space code provided by area LIP [41] and represented by a population code in the distribution of activity in the collicular map [63,64]. Many of the neurons in the SC also have visual receptive fields. Their receptive field center is positioned at the same direction and amplitude as the eye movement that is generated.

At the rostral pole of the map, neurons are clustered that are active during fixation rather than during saccades (Fig. 7B, top panel). In the saccade preparation phase, the activity of these neurons slowly decreases. At the same time, the activity of a subset of the neurons at the map position that represents the saccade target slowly increases (Fig. 7B, middle panel). Because of the slow build-up of their activity in this phase, these neurons are called build-up neurons [65]. When the build-up activity reaches a threshold level, a further subset of SC neurons at that map position becomes activated, the so-called burst neurons (Fig. 7b, bottom panel). These neurons fire a burst of action potentials which triggers the onset of the saccade. The build-up activity reflects many aspects of saccade preparation and target selection. When several targets are available, build-up activity occurs at all associated places in the map. The strength of the build-up activity depends on the probability that the target will be the goal of the saccade [66]. However, when two targets are presented close together, the saccade may be directed to an average of the two target positions, demonstrating that the SC map can perform vector averaging [67].

The target selection and saccade preparation process in the collicular motor map has been modeled in a neural field approach [68,69]. In this approach, the two-

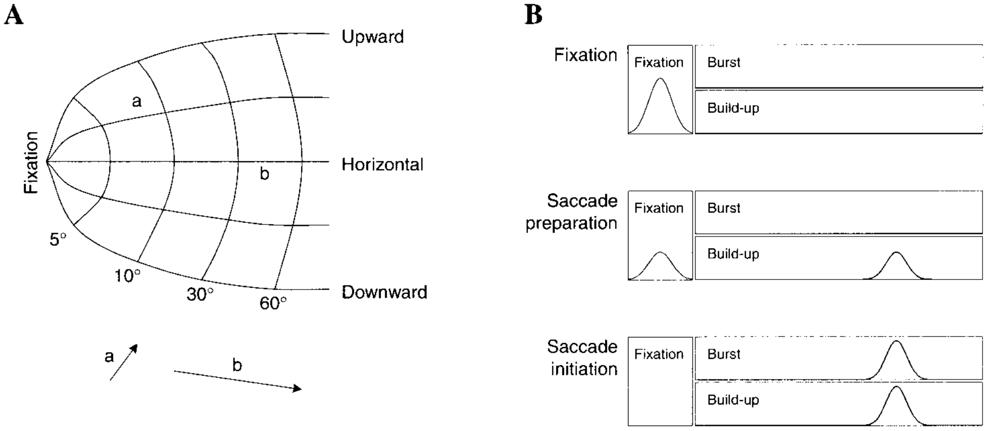

Fig. 7. Saccade initiation in the superior colliculus. (A) The collicular motor map. (B) Distribution of activity in the map during fixation, saccade preparation and saccade initiation.

dimensional map of neurons in the colliculus is treated as a homogeneous excitable field. Lateral interactions are assumed to provide short-distance excitation and long-distance inhibition [70]. In this model, incoming information about the location of saccade targets initiates the build-up of activity at the associated map locations. Simultaneously, the activity in the map is subject to internal dynamics governed by the lateral interactions. The internal dynamics allow to model vector averaging of closeby targets [67,68] and to model influences of multiple targets on saccadic reaction time [69].

When the build-up activity reaches threshold, the collicular burst neurons initiate the saccade. The execution of the saccade is then controlled by the so-called brainstem saccade generator. It consists of several groups of neurons from a number of brainstem nuclei along with the burst neurons and the fixation neurons of the SC. The brainstem saccade generator has to transform spatial target information provided by the saccade initiation system into an appropriate motor program for the eye muscles. Ultimately, this involves the transformation from a spatial map of target position into a temporal signal for the time course of muscle contraction. The brainstem saccade generator functions as a feedback controller (see Section 3.2). It receives a desired gaze displacement as input, which is provided by the SC burst neurons. The output is the signal to the oculomotor neurons. Two other important pathways of the controller are the feedback signal about eye position and the inhibitory pathway that suppresses the fixation neuron activity in order to release fixation. Many models of the brainstem saccade generator fall within this general scheme (e.g. [71–77]). They differ in the exact nature of the input signal and in the way eye position feedback is generated and used.

A saccade aligns the direction of gaze with a target direction in space. The target direction specifies two positional angles of the eye (azimuth and elevation). Yet, the eyeball has three degrees of freedom of movement, the third being torsion around the line of sight. Hence, the saccade target direction does not fully specify the final eye orientation. The saccadic system introduces a further constraint (called Listing's law [11]) that moves the eye such as to minimize torsion [78,79]. This constraint is sensible for two reasons. First, because rotations are non-commutative, a sequence of saccades that does not follow Listing's law would lead to the build-up of strong torsion of the eyeball over time, which would strain the muscles. Second, minimizing torsion ensures that the image of the world always remains in approximately the same orientation in the eye [11,79]. It is still under debate which part of the saccade pathway is responsible for the physiological implementation of Listing's Law [41,80].

Because the duration of a normal saccade is shorter than the latencies of visual input to the brain, the saccadic system cannot receive visual feedback while the saccade is ongoing. Errors in saccade targeting are conveyed to the system only after the saccade is finished. Thus, the entire saccade programming must be based on presaccadic visual information in an open loop sense. It is therefore important for the saccadic system to closely monitor errors in saccade targeting and adjust the saccade programming based upon recent performance. For this reason, saccade gain is plastic [81]. When a subject is instructed to make saccades to a target A that is suddenly moved to another position B during the saccade, initially all saccades miss the target. After about a 100 trials, however, the saccadic system has learned to associate the presaccadic target position A with a saccade that brings the eye to the postsaccadic position B. This requires an adaptation of the gain of the saccade. Such an adaptation is not seen at the level of the superior colliculus [82], suggesting that it occurs in the saccade generator downstream from the colliculus, probably in the cerebellum [81,83,84].

The requirements for the control system and for the input signal become more complicated when combined eye–head movements are considered. In this case, both eye and head movement have to be controlled such that the gaze reaches the target and stays there. Typically, the eye movement is much quicker than the head movement. Therefore, even when both components are initiated at the same time, the eye movement is first to align gaze with the target. Then, as the head follows, the eye has to be counter-rotated to the head in order to keep gaze on the target. This process involves a complicated interaction between the two movement components. It has been proposed that both components might be driven by a common gaze command [85]. But recent experiments suggest that more likely each component receives an independent control signal [86].

5.2. Spatial representations during saccadic gaze shifts

Saccadic eye movements impose problems for the stability of vision. A saccade rapidly and often drastically changes the view of the world that is projected on the retina. Moreover, each saccade induces strong and fast image motion on the retina as it sweeps across the visual image. Typically, however, we are not aware of the retinal image motion generated by a saccade nor of any image displacement after the saccade [87–89]. Both phenomena show that vision is temporally suppressed during saccades. The saccade-induced change of the view of the visual scene, moreover, enforces a match of identical image elements before and after the saccade. Otherwise we would fail to experience a stable environment. Transient changes of the apparent position of a briefly flashed object before a saccade illustrate the mechanisms of transsaccadic visual stability. Just before the beginning of a saccade, the apparent position of briefly displayed objects in the visual scene changes. There is a strong shift in the direction of the saccade anticipating the saccade and compensating for its effects [90,91]. The magnitude of shift varies with position in the visual field, implying a transient compression of the metric of space just before the saccade [92]. However, the compression is less robust than the shift, and is not found under all conditions. It is mainly driven by visual information available in the postsaccadic image [93], while the shift is thought to reflect the efference copy signal [90,91]. The presaccadic position shifts have been linked with the mechanism of receptive field updating in parietal cortex [56].

5.3. Reaching and pointing

Goal-directed movements of the arm are much more difficult to control than eye movements. They involve several joints and hence possess a much higher degree of freedom. Here we will only provide a very brief overview of the main theoretical concepts involved in arm motor control. For more detailed information the reader is referred to a number of review articles [94–98].

The central problem of arm motor control is to specify a unique set of movement parameters in the high-dimensional state space of the joints of the arm. This involves complex control problems for the kinematics and/or dynamics of the joint coordinates. Most approaches to this problem employ optimization schemes. In these schemes, a sequence of joint coordinates is established that minimize some property of the movement. Important examples are movement jerk (i.e. the third derivative of position) [94], torque change [99], or joint stiffness [100]. The equilibrium point hypothesis [101,102] asserts that much of the control of an arm movement is carried out by the passive properties of the muscoloskeletal system. A specific muscle innervation will drive the arm to a certain associated joint configuration, the equilibrium point, which forms a stable attractor of the force field generated by the muscles. In this case, the motor controller only needs to specify the equilibrium point [95,103].

It is also important to observe that normal arm movements do not use all the possibilities that the arm has. Natural arm movements towards a given point in space typically lead to a single posture of the arm. This posture is associated with the position of the target point relative to the shoulder. While many other postures would be possible, the system often uses only a single one. Thus, arm movements normally behave in a more constraint fashion in which only a smaller number of degrees of freedom is actually exploited [94,104]. This bears some analogy to the reduction of the degrees of freedom for saccadic eye movements by Listing's Law (see Section 5.1).

The motor system for arm movements involves areas in the parietal cortex, cortical motor areas, the cerebellum, the basal ganglia, subcortical motor structures, and the spinal cord [1,2]. The motor command neurons in the primary motor cortex are within the highest level of this command hierarchy. These neurons reveal a tuning for the direction of the movement of the hand in space. They are thought to provide a population code for arm movement direction [97,105]. This code can be directly obtained from the distributed encoding of target position in space in the parietal cortex [51]. However, the firing rate of neurons in primary motor cortex is influenced by many other parameters of the movement such as its starting position [106], the orientation of the arm [107], and the load to be moved [108]. Because these parameters influence neuronal firing rate but do not influence the direction of the hand movement, the population may provide more information than just the movement direction. This suggests a more complicated population code which takes proprioceptive feedback about arm position into account [109].

6. Motion

Sensing and interpreting motion are essential for many behavioral tasks. Tracking a moving object with the eye or the hand or controlling one's own motion in the environment requires the determination and analysis of movement in the visual field. Perceiving motion is also helpful for other information processing tasks. Motion can be used to group objects together or to separate objects from their background. The brain contains specialized mechanisms that detect and analyze motion and that transfer information about motion to motor networks involved in motion tasks. This section will first describe how motion is detected and analyzed and later describe how motion information is used in sensorimotor behavior.

6.1. Visual motion detection

An object moving across the visual field induces a changing pattern of illumination on the retina. Motion sensitive neurons respond to spatio-temporal luminance changes when motion is into their individual preferred direction. Such direction-selective responses are already found in the retina, but also in most retinal recipient structures, and in many cortical areas. Many techniques have been proposed to estimate motion from time-varying images (overview and comparison in [110]). Two main classes of models for neuronal motion detection and direction selectivity are correlation models and gradient models.

Correlation models compare the light intensity at one location at a specific time with the light intensity at another location at a later time [111]. The first such models were proposed in the late 1950s and 1960s [112,113]. The basic principle is shown in Fig. 8A. The signal from the first image position is delayed and compared to the signal from the second image position by a coincidence detector, which is modeled as a multiplication of the two signals. This detector responds to a luminance change at the two positions with a specific temporal profile, i.e., to a particular contrast frequency. However, this arrangement alone is not sufficient to truly detect motion in a particular direction. It would also respond to a continuous uniform illumination, because then the time delay does not matter anymore. This ambiguity is re-

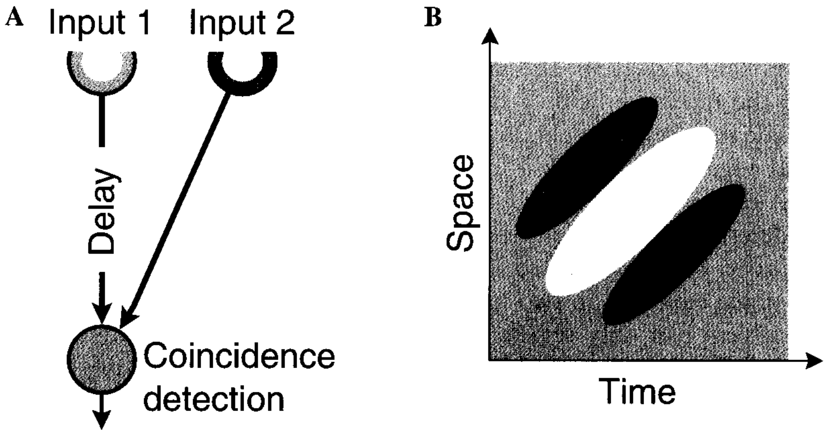

Fig. 8. Motion detection by spatio-temporal correlation. (A) Correlation detector. (B) Spatio-temporal filter.

solved by comparing the outputs of two detectors that are mirror images of each other (opponent detectors) [112,111]. But even then, the detector is direction selective but it cannot be selective for the speed of motion. For instance, two gratings of different spatial frequencies can be moved along the detector at different speeds and yield the same response. Speed information can be gained, however, through the analysis of a population of detectors [114].

A variant of the correlation approach are the motion energy models [115,116] (Fig. 8B). These models use linear spatio-temporal filters that establish time delays for some sub-parts of the receptive field. In the example of Fig. 8B the receptive field consists of an excitatory region (light ellipse) flanked by two inhibitory regions (dark ellipses) in a Gabor-like arrangement. Time delays between these areas can be expressed as an orientation of the receptive field in space-time (Fig. 8B). This receptive field structure establishes a filter that responds preferentially to a spatio-temporal luminance change that is aligned with the long axis of the excitatory region. The outputs of opponent filters are squared and summed to obtain a measure of the total motion strength, or motion energy [115,116]. Subsequent models have refined this general structure, either to make it more consistent with human psychophysical data from motion perception [117,118], or to allow the estimation of speed through population analysis [119]. More recent work has elaborated on these procedures in order to closely resemble the properties of direction selective neurons in the primary visual cortex [120,121].

Gradient models attempt to calculate local velocity from the local spatial and temporal gradients of luminance [122–125]. They are built around the assumption that the total image luminance E is stationary over time, $\mathrm{d}E/\mathrm{d}t = 0$. If this is true, then the temporal and spatial luminance gradients must sum to zero

$$\frac{\partial E}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial E}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial E}{\partial t} = 0.$$

From this equation, image velocity $(\partial x/\partial t, \partial y/\partial t)$ can be computed once ∇E and $\partial E/\partial t$ are known.

An inherent problem in the neural computation of visual motion is the so-called aperture problem [122,123]. The aperture problem occurs if the moving object is larger than the receptive field of the neuron. The neuron only sees the luminance changes inside its receptive field, i.e., through a limited aperture. For a moving one-dimensional edge seen through a limited aperture, only the motion component orthogonal to the edge can be determined. The motion component along the edge cannot be determined because there is no change of luminance along this direction. Mathematically, this is expressed in the fact that the luminance gradient equation above is a single equation in two unknowns. Hence, a neuron that is subject to the aperture problem can only register one component of the two-dimensional visual motion signal. The aperture problem can be overcome by integrating local motion signals from many neurons over a larger spatial region. Several models for such motion integration have been proposed, differing mainly in the way in which the integration is performed. The integration of motion signals may be performed along the edges of a moving object [126], or over a two-dimensional area [123,127]. The latter case attempts to estimate a dense and smooth velocity field across the entire image. However, when motion measurements are spatially averaged, the problem arises that the edges of moving objects form a discontinuity in the velocity field. Averaging motion signals across this discontinuity compromises the estimation of the motion of the object. Moreover, detecting the discontinuity is important to separate the object from its background. In the spatial integration approach, line processes can be incorporated that break up the integration at discontinuities where the local motion signals change abruptly [124]. Alternatively, one may segment the velocity field into coherently moving parts based upon the reliability of individual local motion measurements [128].

In the primate visual system, the integration of local motion signals happens along the pathway from the primary visual cortex (area V1) to the middle temporal (MT) area and the medial superior temporal (MST) area. Neurons in area MT have been shown to overcome the aperture problem [129,130]. Area MT is an area specifically dedicated to the processing of motion. It contains a high proportion of direction-selective neurons [131] and has been linked to behavioral responses to motion stimuli in lesion [132] and microstimulation [133] studies. Area MT contains a topographic map of motion the visual field [134]. Receptive fields in area MT are much larger than those of area V1, reflecting the motion integration that occurs in MT. They range from about $1\,\mathrm{deg}^2$ in the central visual field up to $100\,\mathrm{deg}^2$ in the periphery. Each neuron in the map responds to visual motion at that map position and in its preferred direction. A combination of the responses of several neurons with different preferred directions can provide a population encoding of motion across the visual field [124,128,135].

6.2. *Motion analysis*

Visual motion is used for many purposes. This includes the determination of the three-dimensional structure of objects [136], the control of tracking movements of the eyes [7], and the guidance of self-motion [137]. Each task requires a dedicated analysis of the visual motion signal. This analysis is performed in areas MT and MST, the ventral intraparietal (VIP) area and area 7A in the parietal cortex.

Much useful information is contained in motion parallax [138,139]. Motion parallax is the difference in the apparent motion of two objects that move with the same physical speed but are positioned at different distances from the observer (cf. Fig. 4C). Near objects, in this case, move faster on the retina than objects located further away. This difference in the speed of motion signals from different depths is the basis for estimation of the three-dimensional structure of moving objects [140–143]. Neurons in area MT have an antagonistic substructure in their receptive fields that suggests a role in the estimation of motion parallax. In addition to the part of the receptive field that generates responses to motion in the preferred direction of the neuron, the receptive field of many MT cells contains an area that reduces the response when it is stimulated with motion in that same direction [144,145]. These neurons respond to differences in local motion, i.e., to motion parallax. Their response properties are useful for the estimation of three-dimensional shape [146,147] and for the estimation of self-motion and depth in the visual scene [148,149].

6.3. Visual tracking by smooth pursuit eye movements

Tracking a moving object involves smooth pursuit eye movements. Smooth pursuit eye movements are performed to continuously keep the image of the object on the fovea, i.e., in the area of high-resolution vision. The cortical pathway that generates smooth pursuit is largely separate from the one that drive saccades. Smooth pursuit is generated in a network consisting of the cortical motion areas MT and MST, the frontal eye field (which is also involved in saccades), the pontine nuclei, and certain parts of the cerebellum (overview in [150]). At the level of the pontine nuclei and the cerebellum the pursuit pathway converges to some degree with the pathways for other eye movements.

Smooth pursuit is governed by a feedback control system [151]. It uses visual motion as input and eye speed as the feedback signal. Different models assume different formats for the visual motion input. This difference is best understood by considering the initiation of pursuit to a target that suddenly starts moving. This is the open-loop situation of the feedback system, in which only the input to the controller, not the feedback signal, is available. When the target starts moving, the movement-induced retinal image motion tells the pursuit system the desired speed of the eye. After a certain latency period of about 100 ms the eye is accelerated towards the desired target speed. Meanwhile a saccade is initiated to bring the moving object into the fovea. Later, in the closed-loop situation, smooth pursuit continues to match the eye velocity to the target velocity. But now the retinal image motion of the target is very small because the eye movement stabilizes the target on the fovea. Hence the retinal image motion does not specify the movement of the target anymore and cannot be directly used as the driving signal for the controller. There are two ways to deal with this problem. Target velocity models construct the velocity of the target in space from the sum of the velocity of the eye movement and the remaining retinal motion of the target [151,152]. The target velocity signal is then used as the input to the controller. Retinal motion models, on the other hand, use directly the retinal image motion as input to the controller and integrate it to determine target velocity in space [153].

Neurophysiologically, the input to the pursuit system originates from motion sensitive neurons in areas MT and MST [150,154–156]. Most of these neurons respond to retinal image motion. Hence they can provide the input required by the retinal motion models [156]. However, part of the neurons in area MST appears to encode target velocity in space rather than image motion on the retina [157–159]. Their signal can provide the required input for the target motion model [152].

6.4. Control of self-motion and posture

Movement of the observer himself induces global image motion of the entire visual field (Fig. 9). This pattern of image motion is called optic flow. It serves as a signal to control self-motion and to stabilize posture. Optic flow is used to control body

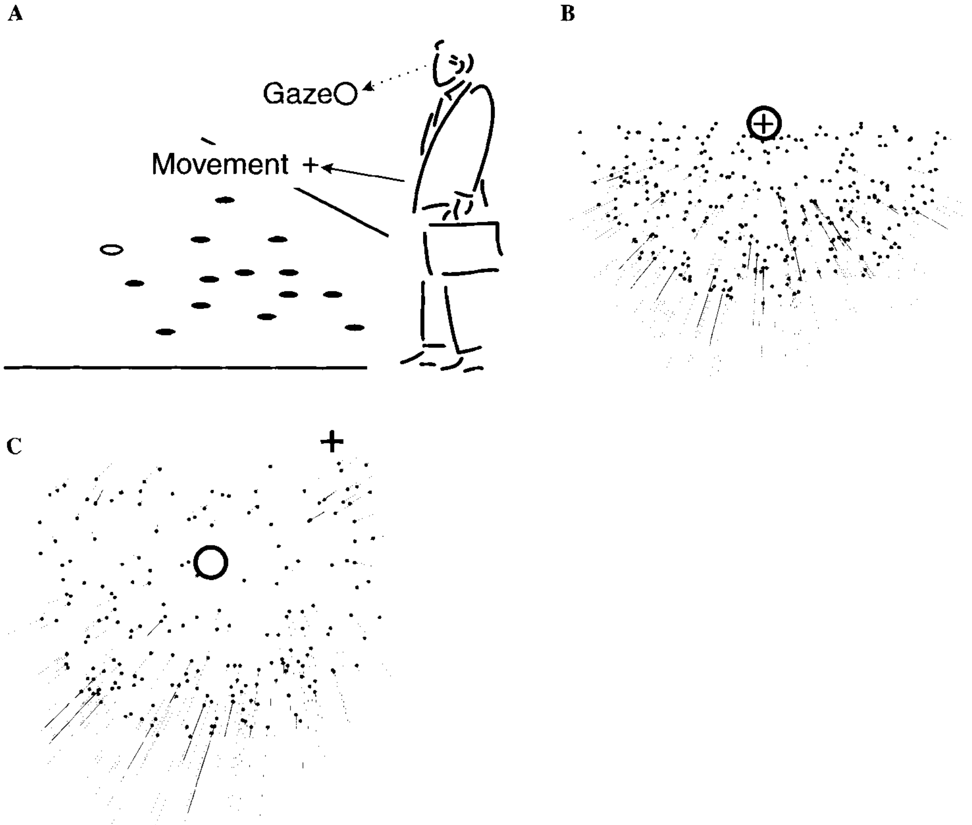

Fig. 9. Optic flow induced by self-motion. (A) An observer moving across a flat horizontal plane. (B) Vector field of image motion induced by forward movement of the observer while he looks directly into the motion direction. (C) Vector field of image motion when the observer performs an eye movement to look at an element on the ground in front of him (circle).

stance [160,161], the speed of self-motion [162], the distance traveled [163], the time-to-collision with obstacles along the path [160,164], and the direction of heading [137].

Much work has concentrated on the last issue, the estimation of heading. The optic flow pattern induced in the eye of a moving observer is determined by the parameters of the movement and the three-dimensional structure of the environment [138]. Mathematically, the problem of inferring the motion of the observer from the pattern of optic flow is ill posed. At any instance, the motion of the eye, like any rigid body motion, can be described by translation and rotation, i.e., with six degrees of freedom. The image motion of an element of the environment depends on these parameters and on its distance from the eye. The structure of the flow pattern is quite simple when only observer translation is considered. In this case, the motion field radiates away from a singular point, the focus of expansion, which is directly equivalent to heading (Fig. 9B). However, when rotational movements of the eye are superimposed on the translation of the observer, the motion field becomes much more complex [165,166] and the singular point is no longer associated with heading (Fig. 9C).

Heading estimation then requires the determination of the direction of translation in the presence of rotational flow disturbances. This can be framed as a problem with many unknown parameters, namely the six degrees of freedom in the self-motion plus the distances of all visible points from the eye. Accurate measurement of the retinal flow provides information to solve this problem by registering the direction and speed of every moving point. This allows the mathematical decomposition of the flow into translational and rotational components and the estimation of heading once more than six moving points are given [167]. Usually many more points are available but their measurements are noisy. In this case, redundant information from more than six points can be used to optimally determine heading [168,169].

A key source of information to separate translational and rotational components of the optic flow is motion parallax. For translational movements of the eye, the induced visual speed of each element is inversely proportional to its distance from the eye. In contrast, a rotation of the eye induces equal angular speed in all image points, independent of distance. This difference is exploited by most neural models for heading estimation from optic flow. Three main classes of models have been proposed. Differential motion models directly capitalize on the properties of motion parallax. By computing differences between adjacent flow vectors they remove the constant rotation component and construct an approximation of the translational component only [170,171,148]. Heading can then be recovered by locating the singular point. The antagonistic receptive field structure of neurons in area MT can provide a starting point for such an analysis [148]. Template-matching models take a different approach. They construct neurons that respond to specific instances of optic flow, i.e., to specific flow patterns. As the number of possible flow patterns in principle is infinite this requires either a very large number of templates [172], or restraining assumptions about the parameters of observer motion or of the environment [173], or a mechanism to approximate an entire set of templates from a few basic templates [174]. A third approach, optimization models, constructs an optimization function that leads to a set of motion parameters that optimally predict the measured flow field. Originally this approach was based on minimizing the squared error between the measured flow field and a possible candidate flow field [168,169]. The parameters that define the best matching candidate flow field were found iteratively. However, the time-consuming iteration can be cut short by geometric considerations, resulting in a fast and robust estimation procedure [175]. This algorithm has been implemented in a neural network [176]. The emerging properties of the elements of this network bear strong resemblance to the properties of neurons in areas MT and MST [177,178]. While the three types of model follow quite different approaches, and make different predictions for the neuronal elements involved, ultimately they all bear on the properties of motion parallax [137]. The first step for this analysis could be provided by receptive field properties in area MT [149]. The optimization model provides a generalization from local to global parallax analysis [149].

The representation of self-motion in areas MT and MST is a good example of the fusion of different sources of information for a common goal. For the separation of translational and rotational (eye-movement) components, motion parallax has been mentioned above as a useful visual cue. A further source of information, however, is provided by eye movement feedback. A nonvisual eye movement feedback signal can be used to adjust the gain of templates in a template model [174] or, more directly, to estimate the eye movement-induced visual motion and to subtract it from the flow pattern in the optimization model [179]. Such an eye movement feedback signal is available in area MST [158], and it is used to obtain heading information in the presence of eye movements [180,181]. In addition, self-motion is also sensed by the vestibular system. Vestibular self-motion signals are integrated with visual self-motion signals in area MST [182,183]. Finally, knowledge of the three-dimensional depth structure of the visual scene can also provide constraints for the evaluation of the flow field [184,185]. Disparity selectivity in area MT contributes to the robustness of the flow representation by a depth-dependent spatial filtering of the flow vectors [135]. It reduces noise among flow vectors with the same motion parallax. Disparity selectivity in area MST may provide a selective weighting of flow signals from distant objects which enhances the separation of translational and rotational components of the flow field [186].

The direct coupling of optic flow to motor output has been investigated in detail for the control of posture. When a standing, stationary subject experiences a low-frequency, low-amplitude expanding and contracting flow pattern, the subject will unconsciously sway back-and-forth along with the pattern movement [160,161]. Hence visual motion is used for postural control, and must be integrated with vestibular and somatosensory signals [187]. The swing and phase coupling behavior has been modeled in a dynamical systems approach. The original idea was that the expansion rate of the flow field can directly drive a passive dynamic system for postural responses [188]. However, detailed comparison with experimental data suggests that the expansion rate of the flow field rather couples into a system that actively generates postural responses [189]. Again, flow field input to this system appears to be provided by area MST [190].

6.5. Gaze stabilization during self-motion

The coupling of perception and action also manifests itself in the oculomotor behaviors during self-motion. In the above section, we have seen that eye movements give rise to rotational components in the optic flow that severely complicate the task of heading detection – and probably many other optic flow-related tasks, as well. Why then do we perform eye movements during self-motion? The reason is that self-motion creates a problem for stable vision as it sets the entire image of the world on the retina in motion. In order to accurately perceive the environment it is desirable to have a clear and stationary visual image. For this reason, several types of compensatory eye movement reflexes are active during self-motion that move the eye so as to keep the central part of the retinal image stable [7,191]. These gaze stabilization mechanisms use vestibular, proprioceptive, and visual signals.

The requirements of gaze stabilization are quite different for rotation and translation movements. For rotations of the head or body, the entire visual scene moves with a single angular velocity. The rotational vestibulo-ocular reflex (rVOR) very directly uses the signal from the vestibular organs to compensates for rotations of the head by rotating the eyes opposite to the head rotation in a feedback control loop [71,191]. The speed of the eyes in the rVOR closely matches the speed of the head movement such that very good image stabilization is achieved. This is particularly true for fast head rotations, e.g. head oscillations in the 2–8 Hz range. For slower head rotations, ocular compensation increasingly relies on the optokinetic reflex (OKR). The optokinetic reflex tries to null retinal image motion by adjusting the eye speed to the speed of the visual motion in a visual feedback control loop. A combination of the optokinetic and vestibulo-ocular reflexes, which is the normal situation during active movement, results in almost complete image stabilization during head rotations. The VOR and the open-loop OKR both have to predict their effects on the visual image and may receive visual feedback signals only with a comparatively long time delay of about 80 ms. For this reason, their performance is constantly monitored through cerebellar feedback loops and subject to fast adaptation mechanisms [10,13,192].

Translations of the head in space also induce vestibularly driven compensatory eye movements, called the translational vestibulo-ocular reflex (tVOR). However, there are two further complications in this case. First, for geometrical reasons the required speed of the eye movement cannot be determined from the head movement alone. The visual motion of a fixated scene element during translation depends on the distance of the element from the observer. Accurate image stabilization in this case must take the geometry of the visual scene into account. If the object is close to the eye the same head movement would induce a much larger visual motion than if the object is further away. Hence to achieve accurate image stabilization the compensatory eye speed depends on the viewing distance [193,194]. Scaling of eye speed with viewing distance also occurs for the ocular following reflex [194]. The second problem is that during forward translation it is physically impossible to stabilize the entire retinal image. Forward motion induces an expanding pattern of optic flow in which points in different parts of the visual field move in different directions. Hence it is only possible to stabilize the part of the visual image at which gaze is directed. In this case, the tVOR varies with viewing direction. Eye movement is rightward when gaze is directed to the right and leftwards when gaze is directed to the left [195]. In addition to the tVOR, optokinetic reactions to radial optic flow fields use visual information to stabilize gaze during forward translation [196,197]. These eye movements follow the direction of motion that is present at the fovea and parafovea, stabilizing the retinal image in a small parafoveal region.

7. An overall view

The brain uses sensory information to control motor behavior. Initially, the incoming information provided by the different sensory systems is encoded in different sensor-specific formats. Typically, this information is either directly or via a first transformation represented in topographic maps. For the guidance of spatially accurate motor action, multiple multisensory representations of space are formed in the parietal cortex. These representations are distributed across populations of neurons. Different task-dependent space coding formats can be specified by the read-out mechanism of the population activity. Different motor tasks, such as the various types of eye movement, are controlled by separate sensorimotor networks. They can mostly be described as feedback control systems. Sensory information acts as the input signal. Feedback is also provided by sensory information and, in addition, by an efference copy of the motor command. The exact types of these signals and their encoding depends on the specific motor task.

Abbreviations

CNS, Central Nervous System
deg, degree
FEF, Frontal eye fields
LIP, Lateral inter parietal
ms, millisecond
MT, Medial temporal
MST, Medial superior temporal
OKR, Optokinetic reflex
s, second
SC, Superior colliculus
V1, primary visual cortex (area 17)
VIP, Ventral inter parietal
VOR, Vestibulo-ocular reflex
2D, two-dimensional

Acknowledgements

I would like to thank several people who have commented on parts of the manuscript: Martin Giese, Klaus-Peter Hoffmann, Dieter Kutz, Lars Lünenburger, Bas Neggers, Peter Thier, Florentin Wörgötter.

References

1. Kandel, E.R., Schwartz, J.H. and Jessell, T.M. (1991) Principles of Neural Science. Elsevier, Amsterdam.
2. Gazzaniga, M.S. (2000) The New Cognitive Neurosciences. MIT Press, Cambridge, MA.
3. Palmer, S.E. (1999) Vision Science: Photons to Phenomenology. MIT Press, Cambridge, MA.
4. Milner, A.D. and Goodale, M.A. (1995) The Visual Brain in Action. Oxford University Press, Oxford.
5. Felleman, D.J. and Van Essen, D.C. (1991) Cerebral Cortex **1**, 1–47.
6. Young, M.P., Scannell, J.W. and Burns, G. (1995) The Analysis of Cortical Connectivity. Springer, Berlin.

7. Carpenter, R.H.S. (1988) Movement of the Eyes, 2nd Edn. Pion, London.
8. Blake, A. and Yuille, A. (1992) Active Vision. MIT Press, Cambridge, MA.
9. Cruse, H. (1996) Neural Networks as Cybernetic Systems. Thieme, Stuttgart, New York.
10. Wolpert, D.M., Miall, R.C. and Kawato, M. (1998) Trends. Cogn. Sci. **2**, 338–347.
11. von Helmholtz, H. (1896) Handbuch der Physiologischen Optik. Leopold Voss, Hamburg.
12. Von Holst, E. and Mittelstaedt, H. (1950) Naturwissenschaften **37**, 464–476.
13. Ito, M. (1998) Trends. Cogn. Sci. **2**, 313–321.
14. Mather, G., Verstraten, F. and Anstis, S. (1998) The Motion Aftereffect. MIT Press.
15. Marr, D. (1982) Vision. Freeman, San Francisco.
16. De Valois, R.L. and De Valois, K.K. (1988) Spatial Vision. Oxford University Press, Oxford.
17. Daugman, J.G. (1980) Vision Res. **20**, 847–856.
18. Jones, J.P. and Palmer, L.A. (1987) J. Neurophysiol. **58**, 1233–1258.
19. Gibson, J.J. (1950) The Perception of the Visual World. Houghton Mifflin, Boston.
20. Cutting, J.E. (1997) Behav. Res. Meth. Instr. Comput. **29**, 27–36.
21. Howard, I.P. and Rogers, B.J. (1996) Binocular Vision and Stereopsis. Oxford University Press, Oxford.
22. Poggio, G. (1995) Cerebral Cortex **5**, 193–204.
23. Fleet, D., Wagner, H. and Heeger, D. (1996) Vision Res. **36**, 1839–1857.
24. Zhu, Y. and Qian, N. (1996) Neural Comp. **8**, 1611–1641.
25. Wagner, H. and Frost, B. (1993) Nature **364**, 796–798.
26. Ohzawa, I., DeAngelis, G. and Freeman, R. (1996) J. Neurophysiol. **75**, 1779–1805.
27. Qian, N. (1994) Neural Comp. **6**, 390–404.
28. Porr, B., Cozzi, A. and Wörgötter, F. (1998) Biol. Cybern. **78**, 329–336.
29. Anzai, A., Ohzawa, I. and Freeman, R.D. (1999) J. Neurophysiol. **82**, 874–890.
30. Qian, N. and Mikaelian, S. (2000) Neural Comp. **12**, 279–292.
31. Lehky, S.R. and Sejnowski, T.J. (1990) J. Neurosci. **10**, 2281–2299.
32. Trotter, Y., Celebrini, S., Stricanne, B., Thorpe, S. and Imbert, M. (1996) J. Neurophysiol. **76**, 2872–2885.
33. Gnadt, J. and Mays, L. (1995) J. Neurophysiol. **73**, 280–297.
34. Pouget, A. and Sejnowski, T.J. (1994) Cerebral Cortex **4**, 314–329.
35. Andersen, R.A., Snyder, L.H., Li, C.-S. and Stricanne, B. (1993) Curr. Opin. Neurobiol. **3**, 171–176.
36. Galletti, C. and Battaglini, P.P. (1989) J. Neurosci. **9**, 1112–1125.
37. Galletti, C., Battaglini, P.P. and Fattori, P. (1995) Eur. J. Neurosci. **7**, 2486–2501.
38. Bremmer, F., Ilg, U.J., Thiele, A., Distler, C. and Hoffmann, K.-P. (1997) J. Neurophysiol. **77**, 944–961.
39. Andersen, R.A., Bracewell, R.M., Barash, S., Gnadt, J.W. and Fogassi, L. (1990) J. Neurosci. **10**, 1176–1196.
40. Boussaoud, D., Jouffrais, C. and Bremmer, F. (1998) J. Neurophysiol. **80**, 1132–1150.
41. Van Opstal, A. and Hepp, K. (1995) Biol. Cybern. **73**, 431–445.
42. Zipser, D. and Andersen, R.A. (1988) Nature **331**, 679–684.
43. Mazzoni, P., Andersen, R.A. and Jordan, M.I. (1991) Proc. Nat. Acad. Sci. USA. **88**, 4433–4437.
44. Goodman, S.J. and Andersen, R.A. (1990) International Journal of Conference on Neural Networks, Vol. II, pp. 381–386, San Diego.
45. Pouget, A., Fisher, S.A. and Sejnowski, T.J. (1993) J. Cog. Neurosci. **5**, 150–161.
46. Bremmer, F., Pouget, A. and Hoffmann, K.-P. (1998) Eur. J. Neurosci. **10**, 153–160.
47. Pouget, A. and Sejnowski, T.J. (1997) J. Cog. Neurosci. **9**, 222–237.
48. Moody, J. and Darken, C. (1989) Neural Comp. **1**, 281–294.
49. Poggio, T. and Girosi, F. (1990) Proc. IEEE **78**, 1481–1497.
50. Salinas, E. and Abbott, L. (1996) Proc. Natl. Acad. Sci. USA. **93**, 11956–11961.
51. Salinas, E. and Abbott, L. (1995) J. Neurosci. **15**, 6461–6474.
52. Duhamel, J.-R., Bremmer, F., Ben Hamed, S. and Graf, W. (1997) Nature **389**, 845–848.
53. Fogassi, L., Gallese, V., di Pellegrino, G., Fadiga, L., Gentilucci, M., Luppino, G., Matelli, M., Pedotti, A. and Rizzolatti, G. (1992) Exp. Brain Res. **89**, 686–690.

54. Boussaoud, D., Barth, T.M. and Wise, S.P. (1993) Exp. Brain Res. **91**, 202–211.
55. Bremmer, F., Distler, C. and Hoffmann, K.-P. (1997) J. Neurophysiol. **77**, 962–977.
56. Duhamel, J.-R., Colby, C.L. and Goldberg, M.E. (1992) Science **255**, 90–92.
57. Umeno, M. and Goldberg, M. (1997) J. Neurophysiol. **78**, 1373–1383.
58. Walker, M., Fitzgibbon, E. and Goldberg, M. (1995) J. Neurophysiol. **73**, 1988–2003.
59. Quaia, C., Optican, L.M. and Goldberg, M. (1998) Neural Networks **11**, 1229–1240.
60. Colby, C. and Goldberg, M. (1999) Ann. Rev. Neurosci. **22**, 319–349.
61. Thier, P. and Andersen, R.A. (1998) J. Neurophysiol. **80**, 1713–1735.
62. Schall, J.D. and Thompson, K.G. (1996) Ann. Rev. Neurosci. **22**, 241–259.
63. Van Gisbergen, J.A.M., Van Opstal, A.J. and Tax, A.A.M. (1987) Neuroscience **21**, 541–555.
64. Lee, C., Rohrer, W.H. and Sparks, D.L. (1988) Nature **332**, 357–360.
65. Munoz, D. and Wurtz, R. (1995) J. Neurophysiol. **73**, 2313–2333.
66. Basso, M.A. and Wurtz, R. (1997) Nature **389**, 66–69.
67. Ottes, F.P., Van Gisbergen, J.A.M. and Eggermont, J.J. (1984) Vision Res. **24**, 1169–1179.
68. Kopecz, K. and Schöner, G. (1995) Biol. Cybern. **73**, 49–60.
69. Kopecz, K. (1995) Vision Res. **35**, 2911–2925.
70. Amari, S.I. (1977) Biol. Cybern. **27**, 77–87.
71. Robinson, D.A. (1981) Ann. Rev. Neurosci. **4**, 463–503.
72. Scudder, C.A. (1988) J. Neurophysiol. **59**, 1455–1475.
73. Waitzman, D.M., Ma, T.P., Optican, L.M. and Wurtz, R.H. (1991) J. Neurophysiol. **66**, 1716–1737.
74. Droulez, J. and Berthoz, A. (1991) Proc. Natl. Acad. Sci. USA **88**, 9653–9657.
75. Dominey, P. and Arbib, M. (1992) Cerebral Cortex **2**, 153–175.
76. Van Opstal, A. and Kappen, H. (1995) Network **4**, 19–38.
77. Moschovakis, A. (1994) Biol. Cybern. **70**, 291–302.
78. Tweed, D. and Vilis, T. (1990) Vision Res. **30**, 111–127.
79. Hepp, K. (1995) Vision Res. **35**, 3237–3241.
80. Raphan, T. (1998) J. Neurophysiol. **79**, 2653–2667.
81. Goldberg, M.E., Musil, S., Fitzgibbon, E., Smith, M. and Olson, C. (1993) in: Role of the Cerebellum and Basal Ganglia in Voluntary Movements, eds N. Mana, I. Hamada and M.R. Delong. pp. 203–211, Elsevier, Amsterdam.
82. Frens, M. and Van Opstal, A. (1997) Brain Res. Bull. **43**, 473–483.
83. Schweighofer, N., Arbib, M. and Dominey, P. (1996) Biol. Cybern. **75**, 19–28.
84. Quaia, C., Lefevre, P. and Optican, L.M. (1999) J. Neurophysiol. **82**, 999–1018.
85. Guitton, D., Munoz, D.P. and L., G.H. (1990) J. Neurophysiol. **64**, 509–531.
86. Goossens, H. and van Opstal, A.J. (1997) Exp. Brain Res. **114**, 542–560.
87. Bridgeman, B., Hendry, D. and Stark, L. (1975) Vision Res. **15**, 719–722.
88. Burr, D.C., Morrone, M.C. and Ross, J. (1994) Nature **371**, 511–513.
89. Ilg, U.J. and Hoffmann, K.-P. (1993) Vision Res. **33**, 211–220.
90. Matin, L. and Pearce, D. (1965) Science **148**, 1485–1488.
91. Honda, H. (1989) Percep. Psychophys. **45**, 162–174.
92. Ross, J., Morrone, M. and Burr, D. (1997) Nature **386**, 598–601.
93. Lappe, M., Awater, H. and Krekelberg, B. (2000) Nature **403**, 892–895.
94. Hogan, N. and Flash, T. (1987) Trends. Neurosci. **10**, 170–174.
95. Mussa-Ivaldi, F.A., Giszter, S.F. and Bizzi, E. (1990) Cold Spring Harbor Symposia on Quant. Biol. **55**, 827–835.
96. Soechting, J. and Flanders, M. (1992) Ann. Rev. Neurosci. **15**, 167–191.
97. Georgopoulos, A. (1995) Trends Neurosci. **18**, 506–510.
98. Jeannerod, M., Arbib, M., Rizzolatti, G. and Sakata, H. (1995) Trends. Neurosci. **18**, 314–320.
99. Uno, Y., Kawato, M. and Suzuki, R. (1989) Biol. Cybern. **61**, 89–101.
100. Hasan, Z. (1986) Biol. Cybern. **53**, 373–382.
101. Feldman, A. (1966) Biophysics **11**, 565–578.
102. Bizzi, E., Accornero, W.C. and Hogan, N. (1984) J. Neurosci. **4**, 2738–2744.
103. Mussa-Ivaldi, F.A. (1992) Biol. Cybern. **67**, 491–500.

104. Gielen, C.C.A.M., Vrijenhoek, E.J., Flash, T. and Neggers, S.F.W. (1997) J. Neurophysiol. **78**, 660–673.
105. Georgopoulos, A.P., Schwartz, A.B. and Kettner, R.E. (1986) Science **233**, 1416–1419.
106. Caminiti, R., Johnson, P. and Urbano, A. (1990) J. Neurosci. **10**, 2039–2058.
107. Scott, S. and Kalaska, J. (1995) J. Neurophysiol. **73**, 2563–2567.
108. Kalaska, J.F., Cohen, D.A., Hyde, M.L. and Prudhomme, M. (1989) J. Neurosci. **9**, 2080–2102.
109. Burnod, Y., Grandguillaume, P., Otto, I., Ferraina, S., Johnson, P.B. and Caminiti, R. (1992) J. Neurosci. **12**, 1435–1453.
110. Barron, J.E., Fleet, D.J. and Beauchemin, S.S. (1994) I.J. Computer Vision **12**, 43–77.
111. Borst, A. and Egelhaaf, M. (1989) Trends. Neurosci. **12**, 29–38.
112. Reichardt, W. (1957) Z. Naturforsch. **12**b, 448–457.
113. Barlow, H.B. and Levick, R.W. (1965) J. Physiol. **173**, 477–504.
114. Snippe, H.P. and Koenderink, J.J. (1994) J. Opt. Soc. Am. A **11**, 1222–1236.
115. Adelson, E.H. and Bergen, J.R. (1985) J. Opt. Soc. Am. A **2**, 284–298.
116. Watson, A.B. and Ahumada, A.J. (1985) J. Opt. Soc. Am. A **2**, 322–341.
117. Van Santen, J.P.H. and Sperling, G. (1985) J. Opt. Soc. Am. A **2**, 300–320.
118. Wilson, H., Ferrera, V.P. and Yo, C. (1992) Vis. Neurosci. **9**, 17–97.
119. Heeger, D.J. (1987) J. Opt. Soc. Am. A **4**, 1455–1471.
120. Heeger, D.J. (1993) J. Neurophysiol. **70**, 1885–1898.
121. Simoncelli, E. and Heeger, D. (1998) Vision Res. **38**, 743–762.
122. Marr, D. and Ullman, S. (1981) Proc. Royal. Soc. London B **211**, 151–180.
123. Horn, B.K.P. and Schunck, B.G. (1981) Art. Intell. **17**, 185–203.
124. Koch, C., Wang, H.T. and Mathur, B. (1989) J. Exp. Biol. **146**, 115–139.
125. Johnston, A., McOwan, P.W. and Buxton, H. (1992) Proc. Royal. Soc. London B **250**, 297–306.
126. Hildreth, E.C. (1984) The Measurement of Visual Motion. MIT, Cambridge, Mass.
127. Yuille, A.L. and Grzywacz, N.M. (1988) Nature **335**, 71–74.
128. Nowlan, S. and Sejnowski, T. (1995) J. Neurosci. **15**, 1195–1214.
129. Movshon, J.A., Adelson, E.H., Gizzi, M.S. and Newsome, W.T. (1985) in: Pattern Recognition Mechanisms, eds C. Chagas, R. Gattass and C. Gross. Springer, Heidelberg.
130. Rodman, H.R. and Albright, T.D. (1989) Exp. Brain Res. **75**, 53–64.
131. Maunsell, J.H.R. and Van Essen, D.C. (1983) J. Neurophysiol. **49**, 1127–1147.
132. Newsome, W.T. and Paré, E.B. (1988) J. Neurosci. **8**, 2201–2211.
133. Salzman, C.D., Britten, K.H. and Newsome, W.T. (1990) Nature **346**, 174–177.
134. Albright, T.D. and Desimone, R. (1987) Exp. Brain Res. **65**, 582–592.
135. Lappe, M. (1996) Neural Comp. **8**, 1449–1461.
136. Andersen, R. and Bradley, D. (1998) Trends. Cogn. Sci. **2**, 222–229.
137. Lappe, M., Bremmer, F. and van den Berg, A.V. (1999) Trends. Cogn. Sci. **3**, 329–336.
138. Koenderink, J.J. (1986) Vision Res. **26**, 161–180.
139. Koenderink, J.J. and van Doorn, A.J. (1992) J. Opt. Soc. Am. A **9**, 530–538.
140. Ullman, S. (1984) Perception **13**, 255–274.
141. Grzywacz, N.M. and Hildreth, E.C. (1987) J. Opt. Soc. Am. A **4**, 503–518.
142. Siegel, R.M. and Andersen, R.A. (1988) Nature **331**, 259–261.
143. Hildreth, E., Ando, H., Andersen, R. and Treue, S. (1995) Vision Res. **35**, 117–137.
144. Allman, J.M., Miezin, F. and McGuinness, E. (1985) Annual Reviews of Neuroscience **8**, 407–430.
145. Raiguel, S., Van Hulle, M., Xiao, D., Marcar, V. and Orban, G. (1995) Eur. J. Neurosci. **7**, 2064–2082.
146. Buracas, G. and Albright, T. (1996) Vision Res. **36**, 869–887.
147. Bradley, D., Chang, G.C. and Andersen, R. (1998) Nature **392**, 714–717.
148. Royden, C.S. (1997) J. Opt. Soc. Am. A **14**, 2128–2143.
149. Lappe, M. (1999) in: Neuronal Processing of Optic Flow, ed M. Lappe. pp. 235–268, Academic Press.
150. Ilg, U.J. (1997) Prog. Brain Res. **53**, 293–329.
151. Robinson, D.A., Gordon, J.L. and Gordon, S.E. (1986) Biol. Cybern. **55**, 43–57.

152. Dicke, P.W. and Thier, P. (1999) Biol. Cybern. **80**, 71–84.
153. Krauzlis, R. and Lisberger, S. (1994) J. Comput. Neurosci. **1**, 265–283.
154. Komatsu, H. and Wurtz, R.H. (1988) J. Neurophysiol. **60**, 580–603.
155. Dürsteler, M.R. and Wurtz, R.H. (1988) J. Neurophysiol. **60**, 940–965.
156. Lisberger, S.G. and Movshon, J.A. (1999) J. Neurosci. **19**, 2224–2246.
157. Sakata, H., Shibutani, H. and Kawano, K. (1983) J. Neurophysiol. **49**, 1364–1380.
158. Erickson, R.G. and Thier, P. (1991) Exp. Brain Res. **86**, 608–616.
159. Ilg, U.J. and Thier, P. (1999) Vision Res. **39**, 2143–2150.
160. Lee, D.N. (1980) Philos. Trans. R. Soc. Lond. B **290**, 169–179.
161. Dijkstra, T., Schöner, G., Giese, M. and Gielen, C. (1994) Biol. Cybern. **71**, 489–501.
162. Prokop, T., Schubert, M. and Berger, W. (1997) Exp. Brain Res. **114**, 63–70.
163. Bremmer, F. and Lappe, M. (1999) Exp. Brain Res. **127**, 33–42.
164. Tresilian, J.R. (1999) Trends. Cogn. Sci. **3**, 301–310.
165. Warren, W.H. and Hannon, D.J. (1990) J. Opt. Soc. Am. A **7**, 160–169.
166. Lappe, M. and Rauschecker, J.P. (1995) Biol. Cybern. **72**, 261–277.
167. Longuet-Higgins, H.C. and Prazdny, K. (1980) Proc. Royal. Soc. London B **208**, 385–397.
168. Bruss, A.R. and Horn, B.K.P. (1983) Comp. Vis. Graph. Image Proc. **21**, 3–20.
169. Koenderink, J.J. and van Doorn, A.J. (1987) Biol. Cybern. **56**, 247–254.
170. Rieger, J.H. and Lawton, D.T. (1985) J. Opt. Soc. Am. A **2**, 354–360.
171. Hildreth, E.C. (1992) Vision Res. **32**, 1177–1192.
172. Perrone, J.A. (1992) J. Opt. Soc. Am. **9**, 177–194.
173. Perrone, J.A. and Stone, L.S. (1994) Vision Res. **34**, 2917–2938.
174. Beintema, J. and van den Berg, A.V. (1998) Vision Res. **38**, 2155–2179.
175. Heeger, D.J. and Jepson, A. (1992) I.J. Computer Vision **7**, 95–117.
176. Lappe, M. and Rauschecker, J.P. (1993) Neural Comp. **5**, 374–391.
177. Lappe, M., Bremmer, F., Pekel, M., Thiele, A. and Hoffmann, K.-P. (1996) J. Neurosci. **16**, 6265–6285.
178. Lappe, M. and Duffy, C. (1999) Eur. J. Neurosci. **7**, 2323–2331.
179. Lappe, M. (1998) Neural Networks **11**, 397–414.
180. Bradley, D., Maxwell, M., Andersen, R., Banks, M.S. and Shenoy, K.V. (1996) Science **273**, 1544–1547.
181. Page, W.K. and Duffy, C.J. (1999) J. Neurophysiol. **81**, 596–610.
182. Duffy, C.J. (1998) J. Neurophysiol. **80**, 1816–1827.
183. Bremmer, F., Kubischik, M., Pekel, M., Lappe, M. and Hoffmann, K.-P. (1999) Ann. N. Y. Acad. Sci. **871**, 272–281.
184. van den Berg, A.V. and Brenner, E. (1994) Nature **371**, 700–702.
185. Grigo, A. and Lappe, M. (1998) Vision Res. **38**, 281–290.
186. Lappe, M. and Grigo, A. (1999) Neural Networks **12**, 1325–1329.
187. Mergner, T. and Rosemeier, T. (1998) Brain Res. Rev. **28**, 118–135.
188. Schöner, G. (1991) Biol. Cybern. **64**, 455–462.
189. Dijkstra, T., Snoeren, P.R. and Gielen, C. (1994) J. Opt. Soc. Am. A **11**, 2184–2196.
190. Duffy, C.J. and Wurtz, R.H. (1995) in: Perception, Memory, and Emotion: Frontier in Neuroscience, eds M. Ono, R. Molotchnikoff and Nishijo. Pergamon Press, Oxford.
191. Miles, F.A. (1998) Eur. J. Neurosci. **10**, 811–822.
192. Du Lac, S., Raymond, J., Sejnowski, T. and Lisberger, S. (1995) Ann. Rev. Neurosci. **18**, 409–441.
193. Schwarz, U. and Miles, F.A. (1991) J. Neurophysiol. **66**, 851–864.
194. Busettini, C., Miles, F., Schwarz, U. and Carl, J. (1994) Exp. Brain Res. **100**, 484–494.
195. Paige, G.D. and Tomko, D.L. (1991) J. Neurophysiol. **65**, 1184–1196.
196. Lappe, M., Pekel, M. and Hoffmann, K.-P. (1998) J. Neurophysiol. **79**, 1461–1480.
197. Niemann, T., Lappe, M., Büscher, A. and Hoffmann, K.-P. (1999) Vision Res. **39**, 1359–1371.