

NEURAL MECHANISMS FOR THE RECOGNITION OF BIOLOGICAL MOVEMENTS

Martin A. Giese^{†} and Tomaso Poggio[‡]*

The visual recognition of complex movements and actions is crucial for the survival of many species. It is important not only for communication and recognition at a distance, but also for the learning of complex motor actions by imitation. Movement recognition has been studied in psychophysical, neurophysiological and imaging experiments, and several cortical areas involved in it have been identified. We use a neurophysiologically plausible and quantitative model as a tool for organizing and making sense of the experimental data, despite their growing size and complexity. We review the main experimental findings and discuss possible neural mechanisms, and show that a learning-based, feedforward model provides a neurophysiologically plausible and consistent summary of many key experimental results.

COGNITIVE NEUROSCIENCE

In recent years the amount and complexity of data in several areas of cognitive neuroscience have increased substantially. As a consequence, pure intuition, and the qualitative mental models associated with it, are becoming less appropriate for interpreting experimental results and for planning new experiments. We believe that quantitative computational theories can be an effective tool for summarizing existing data, and for testing the consistency of possible explanations.

For example, a substantial amount of data about the properties and neural substrates of the recognition of BIOLOGICAL MOVEMENTS is accumulating in neurophysiology, psychophysics and functional imaging, but the underlying computational mechanisms remain largely unknown. It also remains unclear how different experimental evidence is related. Quantitative models might help us to organize our knowledge and to use it to provide explanations and to propose and plan new experiments.

In this article, we review key experimental results relating to the recognition of biological movements. Throughout the article, we use a computational

model as a framework to organize these results. We describe how several experimental results can be accounted for by simple neural mechanisms, under the key assumption that recognition is based on a feedforward architecture that uses learned prototypical patterns. Such prototypes are potentially stored in specific neurons in the ventral and dorsal pathways of the visual system.

Two main sets of questions are addressed. The first is the key to our approach: is it possible to recognize biological movements in a way that is consistent with the experimental data, and that uses plausible neural mechanisms? The second is more specific: what are the roles of the ventral and dorsal pathways for the recognition of biological movement stimuli?

The proposed computational model gives an interpretation of the data that are reviewed and provides answers to these questions. It also points to issues that cannot be answered by the model and by the available experimental results. For instance, how is the information from the two pathways combined, what is the role of time in the ventral pathway, and how does attention influence the recognition process?

**Laboratory for Action Representation and Learning, Department of Cognitive Neurology, University Clinic Tübingen, Spemannstrasse 34, D-72076 Tübingen, Germany.*

†McGovern Institute, Brain Sciences Department and Center for Biological and Computational Learning, Massachusetts Institute of Technology, 45 Carleton Street, Cambridge, Massachusetts 02139, USA. Correspondence to M.G. e-mail: martin.giese@tuebingen.mpg.de doi:10.1038/nrn1057

BIOLOGICAL MOVEMENT

In psychophysics, 'biological motion' often refers specifically to point-light stimuli. We use the term 'biological movements' to characterize stimuli that show the movements of animals and humans, independent of the presentation mode.

SUPERIOR TEMPORAL SULCUS (STS). A sulcus in the temporal lobe of monkey and human cortex that contains areas that are selectively activated by biological movements.

MULTIMODAL

Neurons that respond to sensory input in more than one modality, for example, both visual and auditory stimuli.

Recognition of biological movements

Recognition of complex biological movements — such as locomotion, gestures, facial expressions and motor actions — is biologically important for activities such as detecting predators, selecting prey¹ and courtship behaviour². Gestures and facial expressions are also central to social communication in primates and humans^{3–5}. Humans recognize biological movements accurately and robustly, as was shown in classical psychophysical work by Johansson⁶ (BOX 1).

Neural mechanisms

Only a few electrophysiological experiments have addressed the recognition of biological movements. Some neurons in the SUPERIOR TEMPORAL SULCUS (STS) respond selectively to full-body^{7–11} or hand movements¹². A few of these neurons are view-dependent — their response decreases substantially if the movement is presented from a viewing direction other than the neuron's preferred view^{8–10}. Another set of action-selective neurons has been found in area F5 in monkey premotor cortex^{13,14}. As these neurons also respond when the monkey performs the action, they have been dubbed 'mirror neurons'. Such neurons might be involved in the learning of

motor actions by imitation^{13,15}. Interestingly, mirror neurons and neurons in the STS are often MULTIMODAL^{7,16}.

The neural basis of movement recognition has also been studied using functional imaging in humans^{17–19}. Positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) experiments have found that point-light and natural biological motion stimuli selectively activate areas in the STS^{17–23}. These areas are also selectively activated during observation of mouth and hand movements, and of facial expressions¹⁸. A reported analogue of area F5 in the inferior frontal gyrus in humans shows selective activation during grasping and during the observation and imagery of hand and body movements^{13,17,24,25}. Selective activation for biological movement stimuli has been reported in the occipital and the fusiform face areas¹⁹, and in the lingual gyrus²⁶. Biological movements also activate non-visual areas such as the amygdala and cerebellum^{18,20,22,23}.

The roles of the ventral and dorsal visual processing streams in the recognition of biological movements are unclear. It seems likely that the dorsal pathway, which is specialized for the processing of motion information, contributes substantially to the perception of biological movements²⁷, in particular as the perception of actions is possible without well-defined form information²⁸. At the same time, subjects can recognize gait patterns from individual stationary key frames²⁹, and from stimuli with strongly degraded motion information³⁰, indicating that the ventral pathway is involved. Neurophysiological and imaging experiments support the existence of neurons that respond selectively to human body configurations^{9,19,31}, so biological movements might be recognized by analysing sequences of body shapes that correspond to 'snapshots' from movies of complex movements. fMRI results indicate that normal movement stimuli activate areas in both pathways, whereas point-light stimuli tend not to activate form-selective areas¹⁹.

Learning in biological motion recognition
Learning seems to be important in the recognition of complex movements. Subjects can learn to discriminate between different movement styles³², and between movements of different individuals³³. In addition, learning is fundamental in the recognition of three-dimensional stationary objects³⁴, and the neural representation of objects seems to be based on learned two-dimensional views^{35,36}. This supports the hypothesis that learning is involved in the recognition of complex movements.

Neural model

Before reviewing many of the experimental results on biological movement perception, we describe a model that we use to organize and interpret the data (FIG. 1). The model is based on four assumptions, which are consistent with established anatomical and physiological facts.

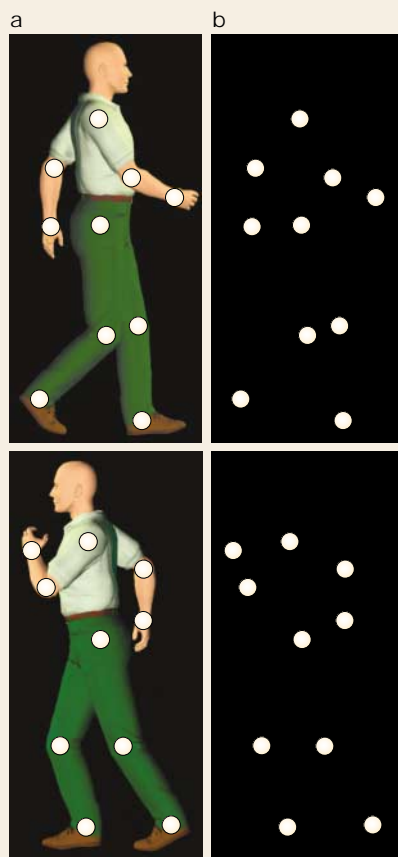
The model is divided into two parallel processing streams^{37–44}, analogous to the ventral and dorsal streams, that are specialized for the analysis of form and OPTIC-FLOW information, respectively.

Both pathways comprise hierarchies of NEURAL FEATURE DETECTORS that extract form or optic-flow features with

Box 1 | Perception of point-light displays

The information carried by biological movements is beautifully illustrated in an experiment devised by the Swedish psychologist Gunnar Johansson⁶. He attached ten light bulbs to the joints of actors (left, a). The actors were videotaped while they performed complex movements, such as walking, running or dancing in the dark. From the videos, which showed only ten light dots moving against a dark background (right, b), subjects could immediately recognize the action. In addition, the dots were spontaneously interpreted as a human being. If the subjects saw individual frames from the videos presented as static pictures, they neither perceived the dots as human, nor were able to identify the actions. (For movies that illustrate the stimuli that have been used to study the recognition of biological movements, see [supplementary information](#))

Subsequent studies have shown that many complex actions can be recognized on the basis of such 'point-light displays', including facial expressions¹³², American Sign Language¹³³, arm movements¹⁰⁰ and various full-body actions⁹⁶. From the same highly impoverished stimuli, subjects could even extract information about subtle details of the action, for example the identity⁹⁸ and gender^{97,102} of walkers, emotional states^{99,100} or the weight of lifted objects¹³⁴. See REF. 135 for a more detailed review of psychophysical studies. (Figure modified, with permission, from REF. 6 © (1973) The Psychonomic Society.)



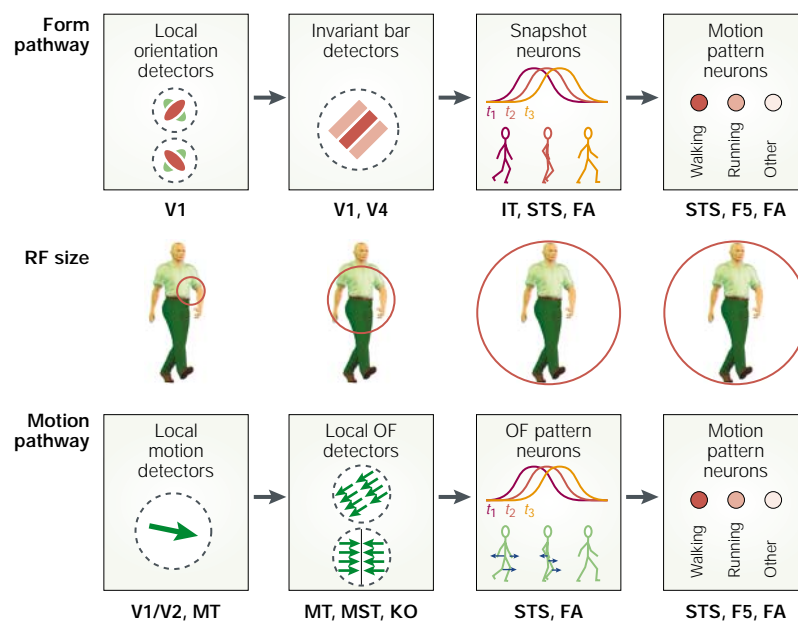


Figure 1 | Hierarchical neural model that provides a unifying interpretation of the data. Overview of the model with two pathways for the processing of form and motion (optic flow). The approximate size of the receptive fields compared to typical stimuli is indicated in the middle row. Insets show the different types of neural detectors at different levels of the hierarchy (see tests referred to in this review). IT, inferotemporal cortex; KO, kinetic occipital cortex; OF, optic flow; RF, receptive field; STS, superior temporal sulcus; V1, primary visual cortex. Other abbreviations indicate corresponding areas in monkey and human visual cortex (see text).

increasing complexity along the hierarchy. The position and size invariance of the feature detectors also increases along the hierarchy.

The model assumes that in its basic, initial operation — akin to ‘immediate recognition’ — the hierarchy in both pathways is predominantly feedforward (apart from local feedback loops), without the need of top–down signals. We do not claim that such signals are not important, in particular for longer stimulus presentations, but we show that without them good recognition performance can be achieved in most cases. This parallels results on the recognition of stationary objects showing that recognition can be achieved with extremely short latencies, making a key role of top–down signals unlikely⁴⁰. Similarly, recordings in the STS have found short latencies for the recognition of biological movements⁴¹. None of this rules out the use of feedback processing. However, it indicates a hierarchical feedforward architecture as the core circuitry underlying ‘immediate’ recognition that might be modulated by recursion and higher-level interactions over longer time intervals.

The representation of motion is based on a set of learned patterns. These patterns are encoded as sequences of ‘snapshots’ of body shapes by neurons in the form pathway, and by sequences of complex optic flow patterns in the motion pathway. This assumption is a central postulate of our model.

Our model extends a previous model for the recognition of stationary objects^{43–45} by integrating form information over time in the ventral pathway, and by

adding the dorsal pathway. The model represents the average belief — often implicit — of several visual physiologists. It is intended to be a quantitative tool to summarize, organize and interpret existing data, and to discuss open questions. Its basic architecture is constrained by neurophysiological data. We have implemented this architecture so that quantitative predictions can be derived through computer simulations.

The architecture of the model (FIG. 1) includes only areas that are primarily visual, and that are involved in movement recognition. The two pathways in the model, which are specialized for the analysis of form and motion (optic flow) information, are kept separate. This is a simplification: in monkey and human brains, the two processing streams interact at several levels^{38,46,47}. Such couplings, for example at the level of the STS, can be easily integrated in the model⁴⁸ and improve the recognition performance without changing the basic results discussed here.

Both pathways consist of a hierarchy of neural feature detectors (TABLE 1). In addition, they contain neural circuits that make recognition sequence-selective. Further details about the model are presented in REF. 49.

Form pathway. The form pathway analyses biological movements by recognizing sequences of ‘snapshots’ of body shapes. Several neurophysiologically plausible models for the recognition of stationary form have been proposed (for example, REFS 43,50,51). The form pathway of our model extends a model for object recognition⁴³ that consists of a hierarchy of neural detectors that process form features of increasing complexity. These detectors correspond to different classes of neurons in the ventral visual pathway. Consistent with neurophysiological data, the receptive field sizes and the position and scale invariance of the neural detectors increase along the hierarchy.

The first level of the form pathway consists of local orientation detectors that model simple cells⁵² in the primary visual cortex (V1). Consistent with other models of simple cells⁵³, these detectors are modelled as GABOR-LIKE FILTERS. The model contains orientation detectors for eight preferred orientations, and two spatial scales that differ by factor 2. The sizes of the receptive fields are in the range of those of neurons in monkey V1 (REF. 54).

The next level of the form pathway contains position- and scale-invariant bar detectors, which extract local orientation information. Within a limited range, their responses are independent of the spatial position and scale of contours within their receptive fields. They might correspond to complex-like cells in area V1 (REF. 52), or to neurons that are increasingly invariant to position changes in areas V2 and V4 (REFS 55,56). The receptive field size of the invariant bar detectors is typical of neurons in area V4 (REF. 57). Many neurons in areas V2 and V4 are selective for more complex form features that are similar to corners or junctions^{55,58}. Such features were not necessary to achieve sufficient selectivity of the model for motion recognition.

A neurophysiologically plausible mechanism for achieving position and scale invariance is the pooling of

OPTIC FLOW

A field of motion vectors that specifies how points of a frame (in an image sequence) are displaced over time. Unlike densely textured scenes, point-light stimuli do not specify a dense, spatially continuous, optic-flow field.

NEURAL FEATURE DETECTORS

Neurons in cortex can often be interpreted as graded ‘detectors’ that are activated when specific features (such as orientation, corners, local motion with defined speed and direction, and faces) are present in their receptive fields.

GABOR-LIKE FILTERS

Gabor functions are defined by sinusoidal functions that are windowed by a gaussian function. They define filters that are localized in the spatial domain as well as in the spatial frequency domain.

Table 1 | Classifications of model neurons

Model neurons	Area	Number of neurons	Receptive field size	References
Form pathway				
Simple cells	V1, V2	1010	0.6°/1.2°	52–54,59
Complex cells	(V1, V2) V4	128	4°	52, 55–58,62,63
(View-tuned) snapshot neurons	IT, EBA, STS, FA	63–840	>8°	19-23,31,36,64–65
Motion pattern neurons	STS, FA, F5	3–40	> 8°	7–15,19–24
Motion pathway				
Local motion detectors	V1, V2, MT	1147	0.9°	54,59,69–70,73–77
Local OF pattern detectors	MT, MST MST, KO	72 (translation) 2 x 50 (expansion/contraction)	3.5°	69,77–78 68,79–87
Complex OF pattern detectors	STS, FA	63–840	> 8°	8–10,19–23
Motion pattern neurons	STS, FA, F5	3–40	> 8°	7–15,19–24

OF, optic flow. Other abbreviations indicate areas in human and monkey visual cortex (see text).

the responses of neurons with similar preferred orientations, but with different receptive field positions and spatial scales^{50,59,60}. We assume that this pooling is accomplished by a nonlinear MAXIMUM-LIKE OPERATION rather than by linear summation^{43,61}. Subpopulations of complex cells in the visual cortex of cats⁶² and neurons in area V4 of macaques⁶³ show behaviour that is compatible with a maximum computation.

The next level of the form pathway contains snapshot neurons that are selective, for instance, for body shapes. These model neurons are similar to view-tuned neurons in monkey inferotemporal cortex (area IT), which are selective for complex shapes^{36,64,65} and can become tuned to complex shapes through learning³⁶. Like view-tuned neurons in area IT⁶⁴, the snapshot neurons have large receptive fields (> 8°) and show substantial position- and scale-invariance. As in previous models of view-tuned neurons^{35,43}, we model the snapshot neurons by GAUSSIAN RADIAL BASIS FUNCTIONS. These neurons receive inputs from the invariant bar detectors on the previous hierarchy level. The centres of the basis functions are adjusted during training so that each snapshot neuron encodes one key frame from a training sequence. In our simulations, each movement pattern is encoded by 21 snapshot neurons representing regularly sampled key frames (this number is not crucial for the performance of the model). The model does not address how an optimum set of key frames can be learned automatically.

Neurons with similar properties in the cortex might be located in area IT in monkeys^{64,65}, and in the STS of monkeys and humans^{10,19–23}. Activity that is selective specifically for human body shapes has been found in the human lateral occipital complex³¹, occipital and fusiform face areas¹⁹ and monkey STS^{9,41}.

The highest hierarchy level of the form pathway consists of motion pattern neurons. These model neurons temporally smooth and summate the activity of all snapshot neurons that contribute to the encoding of the same movement pattern. The temporal smoothing is modelled by a LEAKY INTEGRATOR (see first equation in FIG. 2b). Each motion pattern neuron encodes a single action, such as 'walking' or 'boxing'. Such 'grandmother

cells' are an over-simplification: it is likely that a sparse population code is used, as for the neural encoding of stationary shape^{66,67}. According to physiological data, motion pattern neurons in monkey and human cortex are probably located in the STS^{8–12,19–23}, the premotor cortex (area F5)^{13,25} and possibly the fusiform and occipital face areas¹⁹.

Motion pathway. The motion pathway recognizes biological movements by analysing optic-flow patterns. Consistent with neurophysiological data⁶⁸, it consists of a hierarchy of neural detectors for optic-flow features of increasing complexity. As in the form pathway, the receptive field sizes, invariance of the detectors and complexity of the extracted features increase along the hierarchy.

The first level of the motion pathway consists of local motion detectors that correspond to direction-selective neurons in V1 (REF. 69) and component motion-selective neurons in area MT⁷⁰. Many neurophysiologically plausible models for local motion estimation have been proposed (see, for examples, REFS 71–76). For the simulations reported here, we directly computed optic-flow patterns and calculated the responses of motion-sensitive neurons with physiologically realistic parameters (see REF. 49 for details). Their equivalent receptive field sizes are in the range of direction-selective neurons in V1, and of foveal neurons in area MT^{54,77}.

The second level of the motion pathway consists of neurons with larger receptive fields that analyse the local structure of the optic-flow fields induced by movement stimuli. There are two types of local optic-flow detector. The first is selective for translation flow and corresponds to motion pattern neurons in area MT⁷⁰ with low or bandpass tuning with respect to speed⁷⁸, and has direction tuning curves with a width of about 90° (REF. 69). The model includes neuron populations with four preferred directions and with a receptive field size similar to monkey MT neurons⁷⁷. The second class of local optic-flow detectors is selective for motion edges (horizontal and vertical). Their output signals are computed by combining the responses of two adjacent subfields with opposite

MAXIMUM-LIKE OPERATION
An operation that results in an output signal that approximates the maximum among several input signals. Maximum computation can be approximated by physiologically plausible neural circuits.

GAUSSIAN RADIAL BASIS FUNCTIONS (RBF). Units used to model neurons that are tuned for complex stimuli. Their activation is described by a multi-dimensional gaussian function that depends on the difference between input signals and a constant vector (RBF centre) that defines the input that induces maximal activity.

LEAKY INTEGRATOR
A simplified model for the dynamics of the membrane potential of a neuron. The dynamics are given by a linear differential equation.

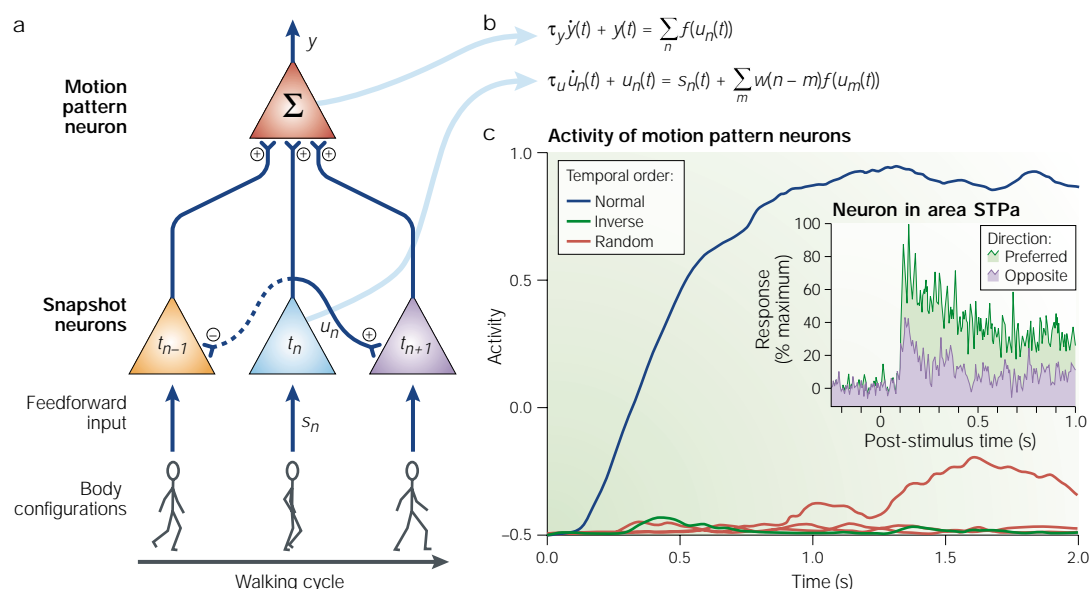


Figure 2 | Neural mechanisms of sequence selectivity. **a** | A model circuit with snapshot neurons and motion pattern neurons that implements sequence selectivity. Each snapshot neuron has asymmetric lateral connections that pre-activate snapshot neurons that encode subsequent body configurations. The other snapshot neurons are inhibited. The output signals of all snapshot neurons that are involved in the encoding of the same motion pattern are summed in a motion pattern neuron. The same circuit is used in the dorsal pathway (with optic-flow pattern neurons instead of snapshot neurons). +, excitation; −, inhibition. **b** | Differential equations describing the dynamics of the snapshot neurons (see REF. 92) and of the motion pattern neurons. Variables: s_n , output signals of radial basis functions trained with learned snapshots; u_n , membrane potential of the snapshot neurons; y , output signal of the motion pattern neuron; w , asymmetric lateral coupling strength; f , sigmoidal nonlinear function; t , time; τ_u , τ_y , time constants (150 ms); m , n , neuron indices. **c** | Inset shows a response of a neuron in area STPa (anterior superior temporal polysensory area) of a macaque monkey that becomes selective with short latency (< 200 ms). Green and purple areas indicate the time course of the neural response for a body stimulus moving in the preferred and anti-preferred direction of the neuron, respectively. The main graph shows the time course of the responses of the motion pattern neurons of the model (form pathway) for stimulus frames presented in the correct temporal order (blue line), reverse temporal order (green line) and random temporal order (red lines). The activity is strongly reduced if the temporal order of the frames is incorrect. Response latency is comparable with the neurophysiological data. Panel c inset modified, with permission, from REF. 41 © (1996) American Physiological Society.

direction preferences (FIG. 1) in a multiplicative way. Neurons with such opponent motion selectivity have been found in several areas in the dorsal processing stream, including areas MT^{79–81}, MSTd^{68,82} and MSTl (REF. 83). They are probably also present in the kinetic occipital area (KO) in humans^{84,85}. Neurons in area MST have substantial position and scale invariance^{86,87}. In our model, such position invariance is obtained by pooling the signals from position-specific motion edge detectors through a maximum operation. The receptive field size of the motion edge-selective detectors is in the range of neurons in areas MT⁵⁴ and MSTl⁸³ in the macaque monkey.

The optic-flow pattern neurons on the third level of the motion pathway are equivalent to the snapshot neurons in the form pathway. Their existence is a prediction of the model. These detectors are selective for complex optic flow patterns that arise for individual moments of biological movement patterns (FIG. 1). Like the snapshot neurons, the motion pattern neurons are modelled by gaussian radial basis functions that receive their inputs from the previous hierarchy level. After training, the centres of the basis functions correspond to the optic-flow patterns that are characteristic for individual moments of the learned movement

sequences (21 neurons per stored pattern). We assume that such optic-flow pattern neurons might be found at different locations in the visual cortex, in particular in the STS^{8–12,19–23}, fusiform and occipital face areas¹⁹, and maybe area MST.

The output signals of the optic-flow pattern neurons are summed and temporally smoothed by the motion pattern neurons of the motion pathway. They are modelled in the same way as the motion pattern neurons of the form pathway. Alternatively, a single set of motion pattern neurons might integrate the information from both pathways (see REF. 48 for an implementation). Motion pattern neurons are probably located in the STS, fusiform and occipital face areas, and maybe area F5 in the premotor cortex.

Sequence selectivity. Movement recognition is selective for temporal order: randomization of the temporal order of the frames of a movie typically destroys the percept of a biological movement, even though the scrambled movie contains the same ‘snapshots’.

In the model, sequence selectivity results from asymmetric lateral connections between the snapshot neurons in the form pathway (and between the optic flow pattern neurons in the motion pathway). With

this circuitry, active snapshot neurons pre-excite neurons that encode temporally subsequent configurations, and inhibit neurons that encode other configurations (FIG. 2a). The network dynamics are given by the equations in FIG. 2b. Significant activity can arise only when the individual snapshot neurons are activated in the 'correct' temporal order. In this case, the network stabilizes a propagating activation pulse with high amplitude, as its feedforward input and the recurrent activation interact in a synergistic way. If the body configurations in the stimulus sequence are in the 'wrong' temporal order, the feedforward and recurrent feedback signals in the snapshot neurons inhibit each other, suppressing the activation (see REF. 88 for a mathematical analysis of the network dynamics). Simulations show that in the model, appropriate lateral connections for the 'correct' sequences can be learned robustly with a simple time-dependent hebbian learning rule^{89,90} from a small number of stimulus repetitions, consistent with psychophysical data⁹¹.

Asymmetric lateral connections are one physiologically plausible implementation of sequence selectivity⁹². In area IT, memory and delay activity⁹³ provide evidence for recurrent connectivity. Sequence selectivity and short-term memory for stationary images might be mediated by similar recurrent neural mechanisms⁹⁴. Alternative mechanisms for sequence selectivity include the classical REICHARDT DETECTOR⁷¹ or equivalent biophysical mechanisms⁹⁵.

FIGURE 2c illustrates the behaviour of the proposed mechanism. The neural activity is strongly reduced if the order of the input frames is reversed or randomized with respect to the training sequence. Otherwise, the activity increases quickly after stimulus onset (in < 200 ms), implying that recognition is possible for stimuli lasting a fraction of a full walking cycle, consistent with psychophysical experiments⁴². The simulation is also consistent with the electrophysiological finding that some STS neurons show selective activity with a latency of < 200 ms⁴¹.

We tested the model with more than 100 video sequences showing natural actions, including different forms of human locomotion, dancing and physical exercises. From these video sequences we generated pixel maps and optic-flow sequences that served as inputs for the two pathways of the model (see REF. 49 for details).

Biological movement perception

In this section, we review several key experimental results. The material is organized according to key properties that characterize the class of neural architectures represented by the model: selectivity, view dependence and robustness. Using the model as an interpretative tool, we also describe data that relate to the roles of the two pathways in motion recognition.

Selectivity and invariance. Humans can easily distinguish many action categories^{6,96}. They can also distinguish between very similar movements, inferring, for instance, the sex, identity or emotional states of moving people^{97–100}. This is the typical trade-off of recognition: it

must generalize across illumination, position, scale and identity of the actor, while being selective enough to distinguish subtle details. This leads to the question of whether the combination of selectivity and generalization that has been observed in humans and in monkeys can be achieved with simple neural mechanisms, and, if so, whether the ventral or the dorsal pathway or both are necessary. This question can be answered only by a quantitative model.

We tested the selectivity and invariance of the two pathways separately in the model (FIG. 3a). The activity of motion pattern neurons trained with 'walking' is shown in the left panels of FIG. 3a for a 'walking' stimulus and 38 distractor patterns (including other forms of locomotion, sports and dancing movements). The motion pattern neurons respond significantly only to 'walking', but not to the distractor patterns. The right panels show the responses of all neurons for all tested action patterns as colour-coded plots, indicating that all motion pattern neurons achieved high pattern selectivity. High selectivity for different action patterns can be achieved with either pathway. Additional simulations⁴⁹ show that each pathway generalizes well over the same action executed several times by the same or different actors.

Consistent with psychophysical data⁹⁸, the selectivity of the model is sufficient to identify people by gait. FIGURE 3b shows the responses of motion pattern neurons in the motion and form pathways, after training with the gaits of two male and two female actors. In both pathways the neurons become activated only for the gait pattern of the person whose gait was presented during training. This result implies that subtle details from biological movement patterns can be recognized using simple neural mechanisms, and by each pathway alone.

The generalization properties of biological movement recognition can be investigated more systematically with stimuli that are generated by MOTION MORPHING. Such morphing stimuli have been used to study action categorization¹⁰¹, gender perception¹⁰² and caricature effects¹⁰³ in movement recognition. The upper panel in FIG. 3c shows a result from a psychophysical experiment¹⁰¹ during which subjects had to categorize motion morphs between four prototypical locomotion patterns — walking, running, limping and marching. The morphs were generated by computing linear combinations of the trajectories of the prototypes using an algorithm based on spatiotemporal correspondences¹⁰⁴. The linear combinations are defined by the equation: morphed pattern = α_1 walking + α_2 running + α_3 limping + α_4 marching, with $0 \leq \alpha_i \leq 1$ and $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = 1$. The weight vectors $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ parameterize a linear space of locomotion patterns. Morphs with different weight combinations had to be categorized into the four categories — 'walking', 'running', 'limping' and 'marching'. FIGURE 3c (left panels) shows the experimental results for the first three categories (for patterns with $\alpha_4 = 0$). The colour indicates the categorization probabilities for seven human subjects. The position of the pixels in the triangles symbolizes the weights of the linear combination $(\alpha_1, \alpha_2, \alpha_3)$. The category judgements vary smoothly with the weights of the prototypes in the linear combination.

REICHARDT DETECTOR

A simple model for local motion detectors, originally studied in the eye of the beetle and the fly. The Reichardt detector multiplies the output signals of two receptors with different positions in the visual field, after delaying or low-pass filtering the output of one of them. The output of the detector is direction-selective. Versions of the Reichardt detector have been used to model direction-selective neurons in the primary visual cortex.

MOTION MORPHING

An algorithm that continuously interpolates between different movement patterns, for example, walking and running. Good morphing algorithms result in interpolations that are very similar to natural movements.

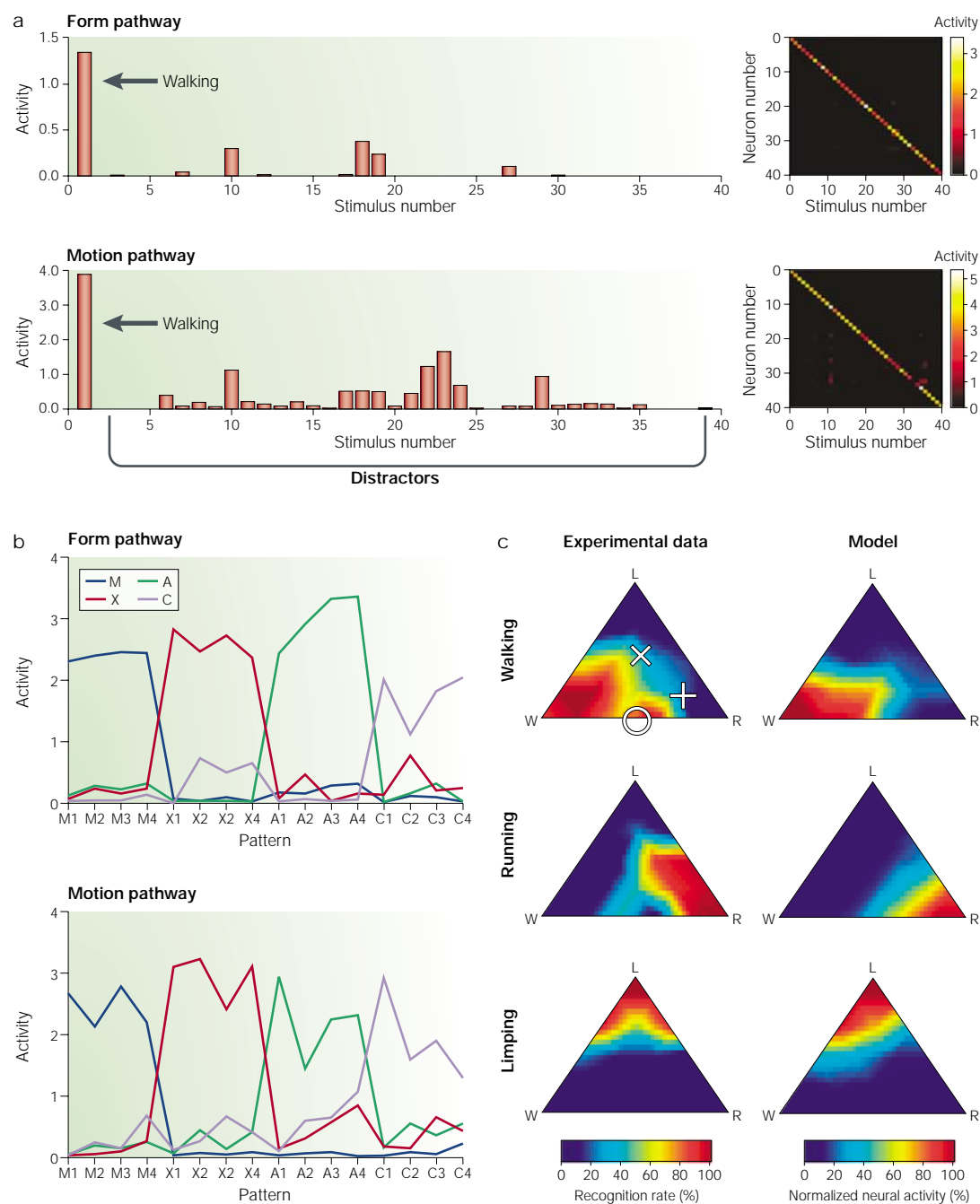


Figure 3 | **Testing selectivity and generalization.** **a** | Left: activation of a motion pattern neurons trained with 'walking' in both pathways for stimulation with 'walking' and 38 distractor patterns. The panels on the right show, as colour-coded plots, the activities of 39 motion pattern neurons, each trained with a different natural action, and each tested with all 39 training patterns. Motion pattern neurons have high pattern selectivity in the form pathway (top panel) and motion pathway (bottom panel). **b** | The model can achieve a degree of selectivity sufficient for person identification by gait. Activities of motion pattern neurons that have been trained with 'walking' executed by different actors (M, X, A and C) are shown for both pathways. Numbers indicate repetitions of the same gait pattern by the same actor. The patterns M1, ..., C1 were used for training. Activities of different motion pattern neurons are indicated by different line styles. **c** | Generalization fields measured with stimuli generated by motion morphing between three prototypes (W, walking; R, running; L, limping). Left: classification rates of the morphs for seven subjects for the categories 'walking', 'running' and 'limping' from a psychophysical experiment. The weight combinations of the prototypes in the morph are encoded by the pixel positions. The corners of the triangles indicate the pure prototypes. Points on the edges of the triangle represent morphs between the two adjacent prototypes, and the points in the triangle represent three-pattern morphs. The distance of the pixel from the corners varies inversely with the contribution of the corresponding prototype. (Examples: O, morph between walking and running with equal weights (50%); X, morph with equal weights (1/3) of all prototypes; +, 20% walking, 20% limping and 60% running.) Right: colour-coded plots of the neural activity of the motion pattern neurons (motion pathway) trained with the three prototypes, and tested with the morphs using the same plotting style. Panel **c** modified, with permission, from REF. 101 © (2002) Elsevier Science.

This indicates that locomotion patterns are embedded into a continuous perceptual motion pattern space. This hypothesis is confirmed by the fact that perceived properties of the morphs also vary smoothly with the weights, and can be accounted for by linear combinations of perceived properties of the prototypes¹⁰¹.

Can such a continuous pattern space of biological movements be adequately represented by the simple hierarchical neural architecture presented in FIG. 1? To address this question, the model was trained with the same four prototypical locomotion patterns and tested with the morphs. The right panels in FIG. 3c show the neural activities of the motion pattern neurons trained with 'walking', 'running' and 'limping' (motion pathway), plotted as for the categorization results. The activities of the motion pattern neurons in the model vary smoothly with the weights of the morph. Even the sizes of the regions in pattern space associated with the individual prototypes in the model coarsely match the psychophysical results. The motion pattern 'walking' occupies a region in pattern space that is larger than the other locomotion patterns for the experimental data and also for the model. As the model treats all training patterns in the same way, this asymmetry must reflect an inherent property of the motion pattern space.

So, the proposed neural model reproduces the generalization properties of biological movement recognition. In other simulations⁴⁹ it also reproduced the following partial invariances of biological movement perception, consistent with experimental data: (1) invariance against position changes of about half of the width of the walker¹⁰⁵; (2) invariance against scale changes up to 1.4 OCTAVES³⁶; (3) invariance against changes of total speed of the walker up to about 2 octaves¹⁰⁶. Non-biological algorithms that are based on alignment with templates or abstract geometrical constraints would predict complete rather than partial invariances.

View-dependence. Another property of biological movement recognition is view-dependence. When a movie of a point-light walker is rotated in the image plane^{107,108} or in depth¹⁰⁹, recognition performance drops substantially. Neural correlates of this effect have been found in the temporo-parieto-occipital area of the macaque monkey⁸ (FIG. 4a, inset). The response of the neuron shown by the green diamond is strongly view-dependent, and decays when the walker rotates in depth.

Are these results consistent with the class of recognition architectures embodied by the model? We would expect them to be, as the representations in both model pathways are based on stored, two-dimensional patterns (snapshots of body configurations and optic-flow patterns). Simulations (white diamond in the inset) reproduce quantitatively the observed view-dependence of the neural responses for rotation. The bar plot in FIG. 3a shows the view-dependence for rotation of the stimulus in the image plane for the response of a motion pattern neuron that was trained with 'walking' for the form pathway. Consistent with psychophysical data¹⁰⁸, the response is degraded substantially for rotations above 30°. The model shows similar effects for the motion pathway.

The same view-dependence has been found in functional imaging experiments¹¹⁰. Biological movement-selective voxels in the STS were determined using an appropriate localizer, and their sum activity was determined for three test stimuli: a point-light walker, an inverted point-light walker (rotated by 180° in the image plane) and a scrambled point-light walker. The measured activity in the STS was much higher for the walker than for the inverted walker (FIG. 5, inset). However, the inverted walker induced more activity than the scrambled walker.

Simulations with our learning-based neural model give the same result (FIG. 5). In the dotted box, the sum activity from the motion pattern neurons of the motion pathway (functionally corresponding to biological motion-selective neurons in the STS) is compared with the experimental data. The inverted walker induces some activity in the model because many of the specific optic-flow features — such as the opponent movements of the feet — are invariant with respect to rotations by 180°. So, view-dependence of the recognition of complex movements is quantitatively consistent with the idea that recognition is based on the matching of learned two-dimensional patterns, whereas mechanisms based on three-dimensional internal models^{111,112} would imply view-independence.

Robustness. Another key property of biological movement recognition is its reliability and robustness: predators must be recognized under difficult conditions, such as clutter or bad illumination. The recognition of point-light walkers illustrates this robustness. It has sometimes been assumed that recognition of such impoverished stimuli requires complex computational mechanisms that determine the link structure of the walker, and that match the movement trajectories of the joints^{113,114}. Can recognition of point-light stimuli be accounted for by the simple neural architecture shown in FIG. 1?

The simulation results presented in FIG. 4b indicate that it can. The figure shows the time-course of the activity of the motion pattern neurons in the form and motion pathways after training with a normal (full-body) 'walking' stimulus. Activities are shown for a normal (full-body) walker, a point-light walker, and distractors (including other locomotion patterns presented as point-light stimuli). In both pathways, the motion pattern neurons become activated for the normal walker stimulus, signalling that they 'recognize' it. In the motion pathway, the point-light stimulus also leads to substantial, but lower activation. The same neuron does not respond to the distractors, showing that the activity is selective for walking. This result is consistent with neurophysiological data from area STPa (anterior superior temporal ploysensory area) of monkeys¹¹⁵ (FIG. 4c). The motion pathway of the model generalizes from full-body to point-light stimuli because the optic-flow field induced by point-light stimuli is a 'sparsely sampled' version of the optic-flow field generated by the full-body stimulus. No such generalization occurs in the form pathway.

OCTAVE
The interval between any two frequencies that have a ratio of 2 to 1.

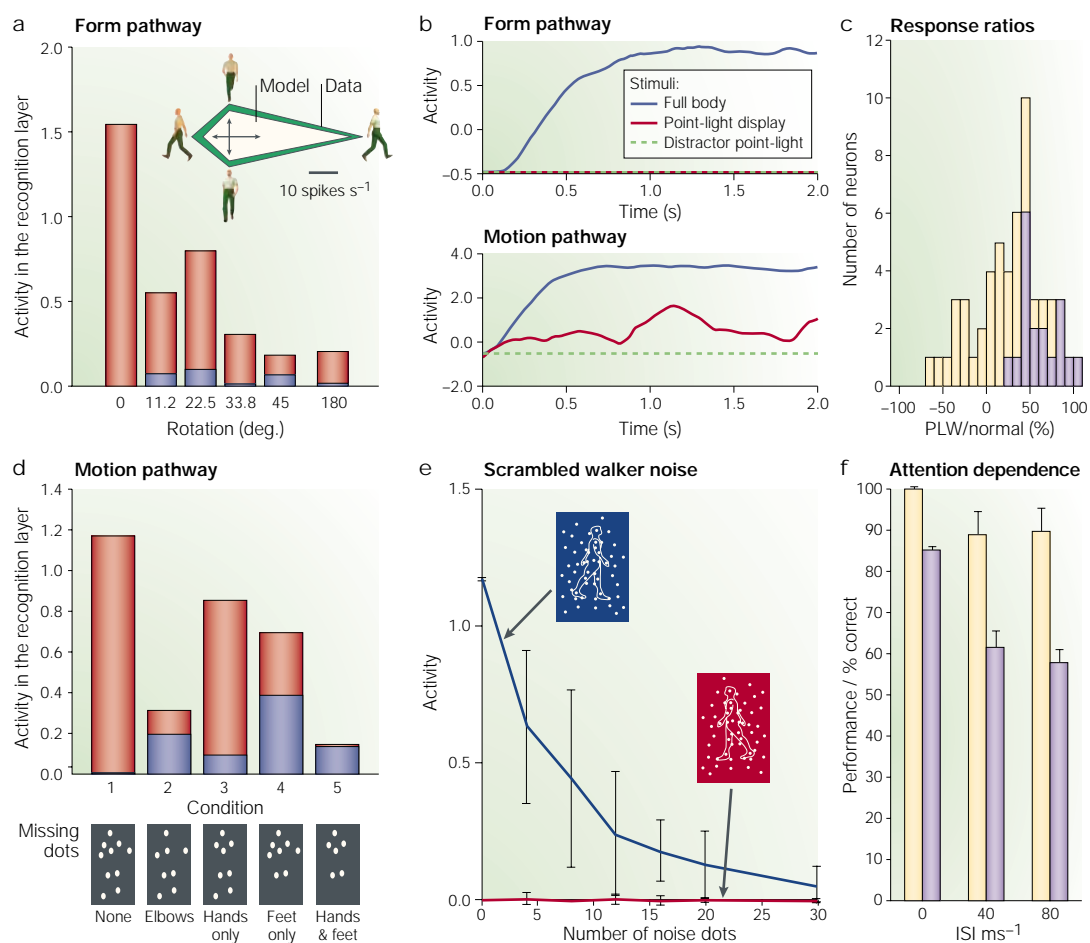


Figure 4 | View variance and robustness. **a** | Activity of the ‘walking’ motion pattern neuron (form pathway) when a ‘walking’ test stimulus is rotated in the image plane against the training view. Red bars indicate the activities of the motion pattern neuron trained with ‘walking’ upright. Blue bars indicate the maximum of the activities of all other motion pattern neurons (encoding running, limping and marching) for the same stimuli. The green diamond shows the activity of a neuron in area TPO (temporo-parieto-occipital region in the superior temporal sulcus) of a macaque monkey as a function of the direction of walking of a walker stimulus. The inner white diamond shows the appropriately scaled response of the ‘walking’ motion pattern neuron in the form pathway of the model for the same stimuli. deg., degrees. Modified, with permission, from REF. 8 © (1985) Elsevier Science. **b** | Activities of the ‘walking’ motion pattern neurons in both pathways for three stimuli: full-body stimulus (blue line), point-light display (red line) and distractor point-light stimuli (red lines). Generalization to point-light stimuli is possible in the motion, but not in the form pathway. **c** | Histogram of the ratios of the response rates of 72 neurons in a macaque monkey area STPa for point-light walkers (PLW) and normal full-body stimuli. Yellow bars indicate the responses of all tested neurons and purple bars the responses of the neurons that were selective for a certain view of a human body and motion direction¹¹⁵. Most selective neurons show reduced responses for point-light stimuli, consistent with the model. Modified, with permission, from REF. 115 © (1994) The International Society for Optical Engineering. **d** | Activities of the motion pattern neurons (motion pathway) for point-light stimuli that were degraded by removing different dots. Red bars indicate the activity of the ‘walking’ neuron and blue bars the maximum of the activities of the other motion pattern neurons. **e** | Activity of motion pattern neuron trained with ‘walking rightwards’, and tested with walking rightwards (blue line) and leftwards (red line) in the presence of different numbers of moving noise dots (see text). The point-light walker consists of 10 dots. Error bars indicate standard deviations over 20 repeated simulations. **f** | Psychophysical evidence for possible dependence of biological movement recognition on attention. For the same masking task, the performance drops substantially if the recognition of the walking direction has to be accomplished simultaneously with a secondary task. Yellow and purple bars show the performance (correct response rate) without and with the secondary task for 66 masking dots. The drop in performance is strongest if the stimulus frames are separated by large inter-stimulus intervals (ISIs), leading to a degradation of low-level motion information. Effects of attention are not included in the model. Modified, with permission, from REF. 119 © (2002) Pion.

Point-light stimuli that are even more degraded can be recognized by the same mechanisms. FIGURE 4d shows the activities of the motion pattern neurons of the motion pathway (trained with walking sequences) for point-light stimuli from which individual dots have been removed. Depending on the missing joints, the recognition performance is more or less degraded. Removing

the elbows and the feet is especially harmful for recognition both in the model and in psychophysical experiments^{27,116}. The effect of the elbows rules out simple explanations based on the maximum speed of the dots. A more detailed analysis shows that opponent motion seems to be crucial for recognition because the elbows specify a characteristic opponent motion feature.

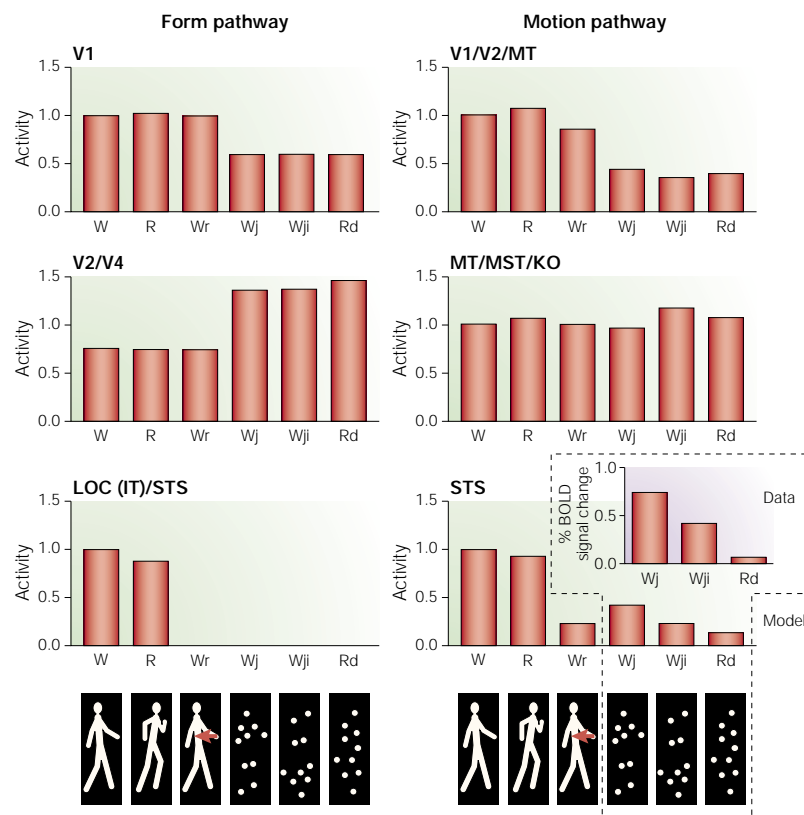


Figure 5 | Prediction of neural activities for a comparison with functional imaging experiments. Bars indicate the sum activities averaged over time for a number of simulated ‘areas’ in both pathways. These areas correspond to the hierarchy levels of the model (see text). Activities are normalized separately for each area (activity for the full-body walking stimulus is set to one). Test stimuli are: W, walking; R, running (full-body); Wr, walking (full-body, reverse temporal order); Wj, walking (point-light stimulus); Wji, point-light walker rotated by 180°; Rd, randomly moving dots (scrambled point-light walker leads to approximately the same results). The dotted box shows the measured blood oxygenation level-dependent (BOLD) signal change (averaged over eight subjects) for stimuli Wj, Wji and Rd from a functional magnetic resonance imaging experiment (data replotted from REF. 110) in comparison with the predicted activity in the motion pattern neurons of the model. Both activities are view-dependent. IT, inferotemporal cortex; KO, kinetic occipital cortex; LOC, lateral occipital complex; MST, medial superior temporal area; MT, middle temporal area; STS, superior temporal sulcus; V1, primary visual cortex.

Masking experiments provide further evidence for the robustness of movement recognition. Many moving background dots can be added to point-light stimuli without greatly impairing recognition^{117,118}, even when the background dots are created by ‘scrambling’ point-light walkers¹¹⁸. Each masking dot has the same movement as one dot of the point-light walker, but a different average position. Surprisingly, our neural model is robust against such masking. FIGURE 4e shows the activity of a motion pattern neuron (in the motion pathway) trained with a full-body stimulus walking rightwards, and tested with point-light stimuli walking rightwards and leftwards, with different numbers of extra, moving, background dots. Even when there are almost three times as many noise dots as dots in the point-light walker, the activation levels for rightwards and leftwards walking are still significantly different. This implies that a right–left discrimination would be possible on the basis of these neural responses in the presence of substantial motion clutter. This robustness results from a hierarchical

architecture that pools neuronal signals using a maximum-like operation and effectively filters out sub-optimal features from the motion clutter⁴³. Humans can further reduce the effect of clutter by top–down ‘attentional’ selection of the region of the walker. Attentional mechanisms are not included in the model. FIGURE 4f shows results from a psychophysical experiment that demonstrates the influence of attention¹¹⁹. When the subjects have to perform a secondary task, the tolerance against masking is strongly reduced.

What are the roles of the two pathways? Our theoretical framework can be used to address the question of the roles of the two pathways in motion recognition. The present data are partially contradictory. Classically, it has been assumed that point-light walkers are analysed primarily in the dorsal (motion) pathway²⁷. This view has been challenged by results on patients with lesions in this pathway who can recognize point-light stimuli^{120,121}. This leads to the question of whether point-light walkers can be recognized by the form pathway alone.

The form pathway of the model does not respond for point-light walkers (FIG. 5), supporting the classical view. This is consistent with the finding from an fMRI experiment that point-light stimuli produce no selective activation in form-selective areas¹⁹. Even when the parameters of the neural detectors are optimized, generalization to point-light stimuli cannot be achieved in the form pathway of the model without losing selectivity for individual action categories. As the first levels of the form pathway respond well for point-light stimuli (FIG. 5), the low responsiveness of the snapshot (output) neurons is not due to inadequate stimulation of the lower level neurons.

A psychophysical experiment that could argue against the involvement of the dorsal pathway in the recognition of point-light displays used point-light stimuli with strongly degraded local motion information³⁰. The lifetime of the dots was limited¹²², and their position was randomly reassigned on the skeleton of the walker in every frame. Detailed simulations (Casile, A. and M.A.G., unpublished observations) show that our model achieves a recognition performance for these stimuli that is comparable with the psychophysical results — exploiting the motion pathway. This is a computational proof that the form pathway is not strictly required for the recognition of such degraded stimuli.

In conclusion, the available data interpreted with our model indicate that both pathways contribute to the recognition of normal biological movement stimuli, and that point-light stimuli are analysed predominantly in the dorsal pathway.

Interpretation of fMRI data. The activity distributions within the two pathways of the model can be compared with fMRI data by assigning different brain areas to the stages of the model. FIGURE 5 shows a possible correspondence that seems compatible with the experimental literature. Such simulations of fMRI results must be interpreted with caution because a homology between

Box 2 | Limitations of the model

The neural architecture we have used in this review is only a first-order approximation with the explicit goal of summarizing, within a consistent quantitative framework, a diverse body of experimental evidence. It provides a plausibility proof that a relatively simple, biologically plausible architecture can account for many properties of the recognition of complex movements. We are aware that the model contains a number of strong simplifications. We list here two of the main shortcomings.

There is experimental evidence for substantial top-down, 'attentional' effects on the recognition of biological motion^{119,128}. Likewise, experimental results demonstrate top-down influences in the two pathways^{38,46,47,129}. The model predicts that attention is not necessarily required for basic recognition of biological motion. Attention is probably needed, however, when motion recognition is too complex — due to, for instance, visual clutter or ambiguous displays — to be solved by one feedforward pass in the form or motion pathway. The model does not yet account for the complexities of every day vision, such as eye movements and attention shifts; it also does not account for the back-projections between cortical areas. It needs to be extended to incorporate top-down mechanisms and their likely anatomical substrates.

In the version of the model described here, we also did not implement biologically plausible mechanisms for the measurement of optical flow from the sequence of real images. There is abundant evidence about the accuracy and robustness of such low-level motion mechanisms, and several good models have been developed^{71–76}. Future implementations of the model will integrate a biologically realistic model for the estimation of motion energy on real video sequences. This will be important for evaluating the limits and accuracy of the proposed architecture for realistic real-world stimuli.

human and monkey cortex is not yet fully established, especially for higher cortical areas, and because of the non-trivial relationship between neural activity and the blood oxygenation level-dependent (BOLD) signal¹²³.

Experimental results show that activity differences that distinguish between biological motion and non-biological motion (scrambled walker) arise only at higher levels of the dorsal pathway^{17–23}, in particular in the STS. This behaviour is expected in the framework of the model (FIG. 5, stimuli Rd versus Wj). Another prediction is that a walker stimulus played in reverse temporal order should elicit significantly less activity in neural populations in the STS that encode forward walking than a normal walker stimulus, whereas only

weak differences are expected at earlier levels of the motion pathway (stimuli W and Wr in FIG. 5). This prediction could be tested in fMRI adaptation experiments¹²⁴ to distinguish neural populations that encode forward and backward walking. Still other predictions can be derived in a similar way from the model and tested in fMRI experiments.

Discussion

Interpretation of existing data. Our review indicates that many existing data on biological movement recognition can be accounted for by a model using established cortical mechanisms. The model's key features are a mainly feedforward architecture and the learning of biological movements by storing prototypical patterns¹²⁵. The proposed model has a number of serious limitations that are discussed in BOX 2. The relationship with other computational methods for motion recognition is briefly discussed in BOX 3.

As the form pathway of our model is an extension of a recently proposed model for the visual recognition of stationary objects^{43–45} our theory includes stationary object recognition as a special case. Many experimental results in this domain could, in principle, be accounted for by the same neural architecture. The form pathway of our model extends these previous models by associating form information over time. Interestingly, time-dependent learning rules^{89,90} of the type that are appropriate for the learning of the sequence selectivity of the snapshot neurons have also been discussed in the context of the learning of view-invariance from image sequences¹²⁶. The learning of view-invariance and of biological movement patterns might therefore share similar neural mechanisms⁹⁴.

Predictions. The available data, interpreted through the lens of the model that we used for this review, motivate several predictions that could be tested in future experiments.

HIDDEN MARKOV MODEL

A finite set of states, each of which is associated with a probability distribution. Transitions among the states are governed by a set of so-called transition probabilities. In a given state, an outcome can be generated according to the associated probability distribution. Only the outcome, not the state, is visible to an external observer and so the states are 'hidden'.

Box 3 | Other models

Several other computational approaches are related to the one used throughout this review. We owe many ideas to previous work; Fukushima's model⁶⁰ is a precursor of the 'standard model' of Riesenhuber and Poggio^{43–45}, which provides the basis of form recognition in our model. Hierarchical feedforward architectures can be traced back to Hubel and Wiesel's original proposal of the simple–complex cell hierarchy⁵⁹, which was later revived by Tanaka in the context of object recognition⁶⁵.

Neural network models for action recognition^{136–138} in computer vision share many aspects with the model of this review. Even though their details are usually not consistent with known properties of the cortex, they provide a computational proof that neural mechanisms similar to the ones discussed in this paper work for complex stimuli and real video sequences.

Other algorithms for motion recognition in computer vision (for reviews, see REFS 139–141) are based on two- or three-dimensional skeleton models that are fitted adaptively, or on optimization under general geometrical constraints^{113,114}. These methods lead to many predictions that are inconsistent with the predictions of our feedforward architecture (view independence, substantial latency after stimulus onset, and so on). Learning-based engineering approaches to motion recognition^{28,139–142} that are based on active shape models or statistical methods (such as HIDDEN MARKOV MODELS) work well in practice, and also predict view dependence. Most of them, however, rely on computational strategies (such as optimization on graphs, dynamic programming) that do not have a direct interpretation in terms of plausible neural circuitries and mechanisms.

For instance, in the dorsal stream, we expect to find neurons tuned to instantaneous optic-flow field patterns that are characteristic for biological movements. In general, neural mechanisms that are similar to the ones that account for the invariant recognition of stationary objects^{34,35,43–45} might also be present in the dorsal pathway for the position- and scale-invariant recognition of complex optic-flow patterns.

We also expect that arbitrary complex movement patterns can be learned, as long as they provide suitable stimulation of the mid- and low-level feature detectors of the two pathways. This is related to the question of how far biological movement recognition is 'special' and potentially innate^{4,127}. We have collected preliminary psychophysical evidence that seems to show that it might not be special⁹¹: synthetic action patterns that do not match any naturally occurring movement can be learned very quickly in the same way as biological movements. In addition, their representations seem to be view dependent.

The two most crucial predictions that follow from the architecture are the existence of neurons tuned to learned biological motions and that attention is not necessary for the basic tasks of motion recognition. If either of these predictions is experimentally falsified, so will be our interpretation of the data. The claim is not that attention does not have any role but rather that, similar to the recognition of stationary objects⁴⁰, attention and top-down influences are not necessary. This of course does not preclude that such effects might be necessary for more sophisticated motion recognition tasks (as illustrated in FIG. 4f).

Data that cannot be readily interpreted. It would be surprising if all available data were to fit our proposed interpretation. A number of important experimental results were not addressed by the interpretation given in this paper. Some of them cannot be explained with the proposed neural model. We did not address the relevance of neurons, for example in the STS¹⁰, that are selective for rigidly translated pictures. Such neurons might be involved in the recognition of translating bodies.

A large and important class of effects that cannot be explained in terms of the proposed feedforward architecture is the influence of attention and top-down effects. FIGURE 4f shows that the human tolerance to motion clutter depends on attention¹¹⁹. Other psychophysical studies, using a visual search experiment, confirm that complex motion recognition tasks require attention¹²⁸. An influence of attention is also indicated by an fMRI experiment showing that a biological motion stimulus did not induce selective activity in the STS during a low-level direction discrimination task, whereas the same area was activated when subjects had to discriminate point-light walkers versus scrambled walkers²³. Evidence for strong top-down influences in motion recognition was obtained in psychophysical experiments during which two human body configurations were presented in sequence with a variable inter-stimulus interval (ISI). For small ISIs,

subjects see motion along the shortest paths between the two body configurations — consistent with the prediction from low-level motion perception. For large ISIs, subjects perceive movement along biomechanically possible paths¹²⁹, indicating an influence of high-level motion recognition on local motion estimation. Similar top-down influences have been reported for depth perception; the depth information indicated by an articulating point-light walker can completely override the stereoscopically specified depth information of the points¹⁰⁹. We believe that the necessary extension of our framework to include all of these top-down effects will be possible by adding back-projection to the basic feedforward architecture, which is the backbone of this review.

Another related question that is not addressed by our interpretation is the role of higher motor areas, and in particular of the mirror neuron system (area F5 in monkeys and BA44 in humans) in the recognition of biological movements. Recent electrophysiological evidence indicates that such neurons might help in understanding partially occluded actions¹³⁰. The responses of many of these neurons, and of some of the action-selective neurons in the STS, are dependent on the presence of target objects for the recognized actions. This is another aspect that we have ignored so far, and that might be integrated in the proposed architecture because the form pathway could be exploited for the recognition of target objects.

Open questions. Our interpretation of the data leaves several open questions: For instance, how exactly is an optimal set of snapshots or optic-flow patterns selected? How is the information from the two pathways combined? Does form or optic flow-based recognition dominate for certain stimulus classes, for example for point-light stimuli as compared to face stimuli? Is there a common area where both pathways converge into a common representation for biological movements? A combination of anatomical and fMRI studies with the investigation of patients with localized lesions should give some insights in this issue. In fact, the effects of localized lesions in the ventral and the dorsal pathway of patients on biological motion recognition can be interpreted directly in terms of the model⁴⁸.

Another question is whether the neural circuits for stationary object recognition and the recognition of snapshots overlap, or if the neural substrates for the representation of moving or articulating objects are localized in separate areas, as indicated by recent fMRI results^{19,31}. This question could be addressed with electrophysiological recordings while stimulating neurons, for example in area IT, with body configurations that are embedded in movement sequences that are compatible or incompatible with biological movements. In psychophysical studies, equivalent priming experiment have yielded results that are compatible with the theoretical framework of this review¹³¹. It would remain a challenge for future electrophysiological and modelling work to unravel the neural correlates of such learning.

1. Ewert, J.-P. Neuroethology of releasing mechanisms: Prey-catching in toads. *Behav. Brain Sci.* **10**, 337–405 (1987).
 2. Morris, D. The reproductive behaviour of the zebra finch (*Poephilla guttata*), with special reference to pseudofemale behaviour. *Behaviour* **6**, 271–322 (1954).
 3. Darwin, C. *The Expressions of the Emotions in Man and Animals* (Appleton, New York, 1872).
 4. Andrew, R. J. Evolution of facial expression. *Science* **142**, 1034–1041 (1963).
 5. Adolphs, R. The neurobiology of social cognition. *Curr. Opin. Neurobiol.* **11**, 231–239 (2001).
 6. Johansson, G. Visual perception of biological motion and a model for its analysis. *Percept. Psychophys.* **14**, 201–211 (1973).
- A classical paper reporting the original experiment in which subjects were able to recognize complex actions from point-light displays that were generated by filming moving actors.**
7. Bruce, C., Desimone, R. & Gross, C. G. Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *J. Neurophysiol.* **46**, 369–384 (1981).
 8. Perrett, D. I. *et al.* Visual analysis of body movements by neurones in the temporal cortex in the macaque monkey: a preliminary report. *Behav. Brain Res.* **16**, 153–170 (1985).
 9. Perrett, D. I. *et al.* in *AI and the Eye* (eds Blake, A. & Troscianko, T.) 181–200 (John Wiley, Chichester, 1990).
 10. Oram, M. W. & Perrett, D. I. Responses of anterior superior temporal polysensory (STPa) neurons to 'biological motion' stimuli. *J. Cogn. Neurosci.* **6**, 99–116 (1994).
- References 8–10 present electrophysiological studies in macaque monkeys with recordings in different parts of the STS during presentation of biological movement stimuli. In different areas, neurons were found that seemed to respond selectively for biological movements of normal and point-light stimuli.**
11. Jellema, T., Baker, C. I., Wicker, B. & Perrett, D. I. Neural representation for the perception of the intentionality of actions. *Brain Cogn.* **44**, 280–302 (2000).
 12. Perrett, D. I. *et al.* Frameworks of analysis for the neural representation of animate objects and actions. Neural representation for the perception of the intentionality of actions. *J. Exp. Biol.* **146**, 87–113 (1989).
 13. Rizzolatti, G., Fogassi, L. & Gallese, V. Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Rev. Neurosci.* **2**, 661–670 (2001).
- A review article about the neural basis of action recognition and imitation with a special focus on the 'mirror neuron' system.**
14. Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. Action recognition in the premotor cortex. *Brain* **119**, 593–609 (1996).
 15. Rizzolatti, G. & Arbib, M. A. Language within our grasp. *Trends Neurosci.* **21**, 188–194 (1998).
 16. Kohler, E. *et al.* Hearing sounds, understanding actions: action representation in mirror neurons. *Science* **297**, 846–848 (2002).
 17. Decety, J. & Grèzes, J. Neural mechanisms subserving the perception of human actions. *Trends Cogn. Sci.* **3**, 172–178 (1999).
- A review article about functional imaging studies of the perception of human actions.**
18. Allison, T., Puce, A. & McCarthy, G. Social perception from visual cues: role of the STS region. *Trends Cogn. Sci.* **4**, 267–278 (2000).
 19. Grossman, E. D. & Blake, R. Brain areas active during visual perception of biological motion. *Neuron* **35**, 1167–1175 (2002).
- An fMRI study that compares activity in different form- and motion-selective areas for biological and non-biological motion.**
20. Bonda, E. *et al.* Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *J. Neurosci.* **16**, 3737–3744 (1996).
- One of the first fMRI studies showing activity that is specific for biological motion in the human STS.**
21. Howard, R. J. *et al.* A direct demonstration of functional specialization within motion-related visual and auditory cortex of the human brain. *Curr. Biol.* **6**, 1015–1019 (1996).
 22. Grossman, E. *et al.* Brain areas involved in perception of biological motion. *J. Cogn. Neurosci.* **12**, 711–720 (2000).
 23. Vaina, L. M. *et al.* Functional neuroanatomy of biological motion perception in humans. *Proc. Natl Acad. Sci. USA* **98**, 11656–11661 (2001).
- An fMRI study that shows an involvement of the opponent motion-selective area KO in biological motion recognition, and an influence of the task on the identified biological motion-sensitive areas.**
24. Iacoboni, M. *et al.* Cortical mechanisms of human imitation. *Science* **286**, 2526–2528 (1999).
 25. Buccino, G. *et al.* Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Eur. J. Neurosci.* **13**, 400–404 (2001).
 26. Servos, P., Osu, R., Santil, A. & Kawato, M. The neural substrates of biological motion perception: an fMRI study. *Cereb. Cortex* **12**, 772–782 (2002).
 27. Mather, G., Radford, K. & West, S. Low-level visual processing of biological motion. *Proc. R. Soc. Lond. B* **249**, 149–155 (1992).
 28. Bobick, A. F. Movement, activity, and action: the role of knowledge in the perception of motion. *Phil. Trans. R. Soc. Lond. B* **352**, 1257–1265 (1997).
 29. Todd, J. T. Perception of gait. *J. Exp. Psychol. Hum. Percept. Perform.* **9**, 31–42 (1993).
 30. Beintema, J. A. & Lappe, M. Perception of biological motion without local image motion. *Poc. Natl Acad. Sci. USA* **99**, 5661–5663 (2002).
 31. Downing, P., Jiang, Y., Shuman, M. & Kanwisher, N. A cortical area selective for visual processing of the human body. *Science* **293**, 2470–2473 (2001).
 32. Pollick, F. E., Fidopiastis, C. & Braden, V. Recognizing the style of spatially exaggerated tennis serves. *Perception* **30**, 323–338 (2001).
- A psychophysical experiment using morphing to generate exaggerations of biological movements. The recognition of spatially exaggerated patterns is improved.**
33. Hill, H. & Johnston, A. Categorizing sex and identity from the biological motion of faces. *Curr. Biol.* **11**, 880–885 (2001).
- A psychophysical study showing that identity can be recognized from facial movements even when identity-specific form information is removed from the face.**
34. Tarr, M. J. & Bühlhoff, H. Image-based object recognition in man, monkey and machine. *Cognition* **67**, 1–20 (1998).
- A review article about view-based object recognition in biological and technical systems.**
35. Poggio, T. & Edelman, S. A network that learns to recognize three-dimensional objects. *Nature* **343**, 263–266 (1990).
 36. Logothetis, N. K., Pauls, J. & Poggio, T. Shape representation in the inferior temporal cortex of monkeys. *Curr. Biol.* **5**, 552–563 (1995).
- An electrophysiological study showing that neurons in area IT of monkey cortex can learn to become selective for artificial three-dimensional shapes. This study also provides evidence for the neural encoding of complex three-dimensional shapes in terms of prototypical views and in particular for size and position invariance being independent of specific experience.**
37. Ungerleider, L. G. & Mishkin, M. in *Analysis of Visual Behavior* (eds Ingle, D. J., Goodale, M. A. & Mansfield, R. J. W.) 549–586 (MIT Press, Cambridge, Massachusetts, 1982).
 38. Felleman, D. J. & van Essen, D. C. Distributed hierarchical processing in the primate visual cortex. *Cereb. Cortex* **1**, 1–47 (1991).
 39. Milner, A. D. & Goodale, M. A. Visual pathways to perception and action. *Prog. Brain Res.* **95**, 317–337 (1993).
 40. Thorpe, S., Fize, D. & Marlot, C. Speed of processing in the human visual system. *Nature* **381**, 520–522 (1996).
- An electroencephalography study showing that selective cortical activity in a categorization task can be detected as early as 150 ms after stimulus onset. This makes it unlikely that complex feedback circuits are needed for basic, 'immediate' categorization tasks.**
41. Oram, M. W. & Perrett, D. I. Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey. *J. Neurophysiol.* **76**, 109–129 (1996).
 42. Johansson, G. Spatio-temporal differentiation and integration in visual motion perception. *Psychol. Res.* **38**, 379–393 (1976).
 43. Riesenhuber, M. & Poggio, T. Hierarchical models of object recognition. *Nature Neurosci.* **2**, 1019–1025 (1999).
- A neurophysiologically plausible model for the recognition of stationary objects that provides the skeleton of the form pathway of our model for motion recognition. The study shows that pooling by a maximum-like operation provides feature selectivity simultaneously with sufficient position and size invariance.**
44. Riesenhuber, M. & Poggio, T. Models of object recognition. *Nature Neurosci.* **3**, Suppl. 1199–1204 (2000).
 45. Riesenhuber, M. & Poggio, T. Neural mechanisms of object recognition. *Curr. Opin. Neurobiol.* **12**, 162–168 (2002).
 46. Kourtzi, S. & Kanwisher, N. Activation in human MT/MST by static images with implied motion. *J. Cogn. Neurosci.* **12**, 48–55 (2000).
 47. Saleem, K. S., Suzuki, W., Tanaka, K. & Hashikawa, T. Connections between anterior inferotemporal cortex and superior temporal sulcus regions in the macaque monkey. *J. Neurosci.* **20**, 5083–5101 (2000).
- An anatomical study providing evidence for strong connections between IT and STS.**
48. Giese, M. A. & Vaina, L. M. Pathways in the analysis of biological motion: computational model and fMRI results. *Perception* **30**, Suppl. 116 (2001).
 49. Giese, M. A. & Poggio, T. Neural mechanisms for the recognition of biological movements and actions. *AI Memo 2002-012/CBCL Paper 219* (Massachusetts Institute of Technology, Cambridge, Massachusetts, 2002).
 50. Perrett, D. I. & Oram, M. W. Neurophysiology of shape processing. *Image Vis. Comput.* **11**, 317–333 (1993).
 51. Rolls, E. T. & Milward, T. A model of invariant object recognition in the visual system: learning rules, activation functions, lateral inhibition, and information-based performance measures. *Neural Comput.* **12**, 2547–2572 (2000).
 52. Hubel, D. H. & Wiesel, T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol. (Lond.)* **160**, 106–154 (1962).
 53. Jones, J. P. & Palmer, L. A. An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *J. Neurophys.* **58**, 1233–1258 (1987).
 54. Dow, B. M., Snyder, A. Z., Vautin, R. G. & Bauer, R. Magnification factor and receptive field size in foveal striate cortex of the monkey. *Exp. Brain Res.* **44**, 213–228 (1981).
 55. Hegdè, J. & van Essen, D. C. Selectivity for complex shapes in primate visual area V2. *J. Neurosci.* **20**, RC61 (2000).
 56. Gallant, J. L., Braun, J. & van Essen, D. C. Selectivity for polar, hyperbolic, and Cartesian gratings in macaque visual cortex. *Science* **259**, 100–103 (1993).
 57. Plonk, M. C., Gattass, R. & Sousa, A. P. Area V4 in Cebus monkey: extent and visuotopic organization. *Cereb. Cortex* **8**, 685–701 (1998).
 58. Pasupathy, A. & Connor, C. E. Response to contour features in macaque area V4. *J. Neurophysiol.* **82**, 2490–2502 (1999).
 59. Hubel, D. & Wiesel, T. Ferrier lecture. Functional architecture of the macaque monkey visual cortex. *Proc. R. Soc. Lond. B* **198**, 1–59 (1977).
 60. Fukushima, K. Neocognitron: a self-organizing neural network model for a mechanisms of pattern recognition unaffected by shift in position. *Biol. Cybern.* **36**, 193–202 (1980).
- A classical neural network model for object recognition that uses a hierarchy of neural detectors in combination with pooling through a maximum operation to account for invariant recognition.**
61. Mel, B. W. & Fiesler, J. Minimizing binding errors using learned conjunctive features. *Neural Comp.* **12**, 731–762 (2000).
 62. Lampl, I., Riesenhuber, M., Poggio, T. & Ferster, D. The max operation in cells in the cat visual cortex. *Soc. Neurosci. Abstr.* **30**, 619 (2001).
 63. Gawne, T. J. & Martin, J. Response of primate visual cortical V4 neurons to two simultaneously presented stimuli. *J. Neurophysiol.* **88**, 1128–1135 (2002).
- An electrophysiological study showing evidence for maximum computation in some neurons in area V4 of monkeys.**
64. Logothetis, N. K. & Sheinberg, D. L. Visual object recognition. *Annu. Rev. Neurosci.* **19**, 577–621 (1996).
 65. Tanaka, K. Inferotemporal cortex and object vision. *Annu. Rev. Neurosci.* **19**, 109–139 (1996).
 66. Young, M. & Yamane, S. Sparse population coding of faces in the inferotemporal cortex. *Science* **256**, 1327–1331 (1992).
 67. Rolls, E. T. & Tovee, M. J. Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex. *Exp. Brain Res.* **73**, 713–726 (1995).
 68. Saito, H. in *Brain Mechanisms of Perception and Memory* (eds Ono, T., Squire, L. R., Raichle, M. E., Perrett, D. I. & Fukuda, M.) 121–140 (Oxford Univ. Press, Oxford, UK, 1993).
- A review article about the dorsal pathway indicating that the analysis of complex optical flow patterns is accomplished by a hierarchy of motion-selective areas, similar to form analysis in the ventral stream.**
69. Smith, A. T. & Snowden, R. J. *Visual Detection of Motion* (Academic, London, UK, 1994).
 70. Rodman, H. R. & Albright, T. D. Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT). *Exp. Brain Res.* **75**, 53–64 (1989).
 71. Poggio, T. & Reichardt, W. Visual control of orientation behavior in the fly. Part II: towards the underlying neural interactions. *Quart. Rev. Biophys.* **9**, 377–438 (1976).
 72. Heeger, D. J. Optical flow using spatiotemporal filters. *Int. J. Comput. Vis.* **1**, 279–302 (1988).
 73. Grossberg, S., Mingolla, E. & Viswanathan, L. Neural dynamics of motion integration and segmentation within and across apertures. *Vision Res.* **41**, 2521–2553 (2001).
 74. Sereno, M. E. *Neural Computation of Pattern Motion* (MIT Press, Cambridge, Massachusetts, 1993).
 75. Nowlan, S. J. & Sejnowski, T. J. A selection model for motion processing in area MT of primates. *J. Neurosci.* **15**, 1195–1214 (1995).
 76. Simoncelli, E. P. & Heeger, D. J. A model of neuronal responses in visual area MT. *Vision Res.* **38**, 743–761 (1998).
 77. Albright, T. D. & Desimone, R. Local precision of visuotopic organization in the middle temporal area (MT) of the macaque. *Exp. Brain Res.* **65**, 582–592 (1987).

78. Lagae, L., Raiguel, S. & Orban, G. A. Speed and direction selectivity of macaque middle temporal neurons. *J. Neurophysiol.* **69**, 19–39 (1993).
79. Allman, J., Miezin, F., & McGuinness, E. Direction- and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT). *Perception* **14**, 105–126 (1985).
80. Xiao, D. K. *et al.* Spatial heterogeneity of inhibitory surrounds in the middle temporal visual area. *Proc. Natl Acad. Sci. USA* **92**, 11303–11306 (1995).
81. Born, R. T. Center-surround interactions in the middle temporal visual area of the owl monkey. *J. Neurophysiol.* **84**, 2658–2669 (2000).
82. Tanaka, K., Fukuda, Y. & Saito, H. Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. *J. Neurophysiol.* **62**, 626–641 (1989).
83. Eifuku, S. & Wurtz, R. H. Response to motion in extrastriate area MSTl: center-surround interactions. *J. Neurophysiol.* **80**, 282–296 (1998).
84. Orban, G. *et al.* A motion area in human visual cortex. *Proc. Natl Acad. Sci. USA* **92**, 993–997 (1995).
85. Dupont, P. *et al.* The kinetic occipital region in human visual cortex. *Cereb. Cortex* **7**, 283–292 (1997).
- References 84 and 85 are fMRI studies describing an area in the occipital cortex — KO — that responds particularly strongly for motion discontinuities.**
86. Graziano, M. A., Andersen, R. A. & Snowden, R. J. Tuning of MST neurons to spiral motions. *J. Neurosci.* **14**, 54–67 (1994).
87. Geesaman, B. J. & Andersen, R. A. The analysis of complex motion patterns by form/cue invariant MSTd neurons. *J. Neurosci.* **16**, 4716–4732 (1996).
88. Xie, X. & Giese, M. A. Nonlinear dynamics of direction-selective nonlinear neural media. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **65**, 051904 (2002).
89. Markram, H., Lübke, J., Frotscher, M. & Sakmann, B. Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* **275**, 213–215 (1997).
- A study showing that synaptic changes depend on the precise timing of pre- and postsynaptic activity. Spike-time-dependent plasticity is a biologically plausible mechanism for learning temporal order.**
90. Sejnowski, T. J. The book of Hebb. *Neuron* **24**, 773–776 (1999).
91. Giese, M. A., Jastorff, J. & Kourtzi, Z. Learning of the discrimination of artificial complex biological motion. *Perception* **31**, Suppl. 117 (2002).
92. Mineiro, P. & Zipser, D. Analysis of direction selectivity arising from recurrent cortical interactions. *Neural Comput.* **10**, 353–371 (1998).
93. Miyashita, Y. & Chang, H. S. Neural correlate of pictorial short-term memory in the primate temporal cortex. *Nature* **331**, 68–70 (1988).
94. Stryker, M. P. Neurobiology. Temporal associations. *Nature* **354**, 108–109 (1991).
95. Koch, C., Poggio, T. & Torre, V. Nonlinear interactions in a dendritic tree: localization, timing, and role in information processing. *Proc. Natl Acad. Sci. USA* **80**, 2799–2802 (1983).
96. Dittrich, W. H. Action categories and the perception of biological motion. *Perception* **22**, 15–22 (1993).
97. Kozlowski, L. T. & Cutting, J. E. Recognizing the sex of a walker from a dynamic point light display. *Percept. Psychophys.* **21**, 575–580 (1977).
- A classical study showing that subjects can recognize gender from point-light walkers.**
98. Cutting, J. E. & Kozlowski, L. T. Recognizing friends by their walk: gait perception without familiarity cues. *Bull. Psychon. Soc.* **9**, 353–356 (1977).
- A classical study showing that subjects can recognize identity from point-light walkers.**
99. Dittrich, W. H., Troscianko, T., Lea, S. E. & Morgan, D. Perception of emotion from dynamic point-light displays represented in dance. *Perception* **25**, 727–738 (1996).
100. Pollick, F. E., Paterson, H. M., Bruderlin, A. & Sanford, A. J. Perceiving affect from arm movement. *Cognition* **82**, B51–B61 (2001).
101. Giese, M. A. & Lappe, M. Measuring generalization fields for the recognition of biological motion. *Vision Res.* **42**, 1847–1858 (2002).
- A psychophysical study of the categorization of biological motion using linear spaces of movement patterns defined by motion morphing.**
102. Troje, N. Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *J. Vis.* **2**, 371–387 (2002).
103. Hill, H. & Pollick, F. E. Exaggerating temporal differences enhances recognition of individuals from point light displays. *Psychol. Sci.* **11**, 223–228 (2000).
- A psychophysical study showing that, by motion morphing, temporal caricatures of arm movement patterns can be generated that improve the recognition of identity from point-light walkers.**
104. Giese, M. A. & Poggio, T. Morphable models for the analysis and synthesis of complex motion patterns. *Int. J. Comp. Vis.* **38**, 59–73 (2000).
105. Verfaillie, K., de Troy, A. & van Rensbergen, J. Transsaccadic integration of biological motion. *J. Exp. Psychol. Learn. Mem. Cogn.* **20**, 649–670 (1994).
106. Pavlova, M. Biological motion perception under various presentation rates. *Perception* **24**, Suppl. 112 (1995).
107. Sumi, S. Upside-down presentation of the Johansson moving light-spot pattern. *Perception* **13**, 283–302 (1984).
- A psychophysical experiment showing impaired biological motion recognition for upside-down stimuli.**
108. Pavlova, M. & Sokolov, S. Orientation specificity in biological motion perception. *Percept. Psychophys.* **62**, 889–899 (2000).
109. Bülthoff, I., Bülthoff, H. H. & Sinha, P. Top-down influences on stereoscopic depth-perception. *Nature Neurosci.* **1**, 254–257 (1998).
- A psychophysical study showing that biological motion recognition is view dependent, and that the depth information indicated by the structure of the walker dominates over the stereoscopically specified depth of the walker dots.**
110. Grossman, E. D. & Blake, R. Brain activity evoked by inverted and imagined biological motion. *Vision Res.* **41**, 1475–1482 (2001).
111. Marr, D. & Nishihara, H. K. Representation and recognition of the spatial organization of three-dimensional shapes. *Proc. R. Soc. Lond. B* **200**, 269–294 (1978).
- A paper on the recognition of articulated figures in terms of structural three-dimensional models.**
112. Hogg, D. Model-based vision: a program to see a walking person. *Image Vis. Comp.* **1**, 5–19 (1983).
113. Hoffman, D. D. & Flinchbaugh, B. E. The interpretation of biological motion. *Biol. Cybern.* **42**, 195–204 (1982).
114. Webb, J. A. & Aggarwal, J. K. Structure from motion of rigid and jointed objects. *Artif. Intell.* **19**, 107–130 (1982).
- A classical paper on the estimation of the structure of articulated objects based on geometrical constraints.**
115. Oram, M. W. & Perrett, D. I. Neural processing of biological motion in the macaque temporal cortex. In *Computational Vision Based on Neurobiology* (ed. Lawton, T. B.) *SPIE Proc.* **2054**, 155–165 (1994).
116. Pinto, J. & Shiffrar, M. Subconfigurations of the human form in the perception of biological motion displays. *Acta Psychol.* **102**, 293–318 (1999).
117. Cutting, J. E., Moore, C. & Morrison, R. Masking the motions of human gait. *Percept. Psychophys.* **44**, 339–347 (1988).
- A psychophysical experiment showing that point-light walkers can be recognized even in the presence of a substantial number of moving background dots and scrambled walker noise.**
118. Bertenthal, B. I. J. & Pinto, J. Global processing of biological motion. *Psychol. Sci.* **5**, 221–225 (1994).
119. Thornton, I. M., Rensink, R. A. & Shiffrar, M. Active versus passive processing of biological motion. *Perception* **31**, 837–853 (2002).
- A psychophysical study using a dual-task experiment to show that the tolerance of biological motion recognition against masking depends on attention.**
120. McLeod, P., Dittrich, W., Driver, J., Perrett, D. & Zihl, J. Preserved and impaired detection of structure from motion by a ‘motion-blind’ patient. *Vis. Cogn.* **3**, 363–391 (1996).
121. Vaina, L. M., Lemay, M., Bienfang, D. C., Choi, A. Y. & Nakayama, K. ‘Intact’ biological motion and ‘structure from motion’ perception in a patient with impaired motion mechanisms: a case study. *Vis. Neurosci.* **5**, 353–369 (1990).
- References 120 and 121 describe studies of patients that have lesions in the dorsal pathway and that can still recognize biological motion.**
122. Neří, P., Morrone, C. & Burr, D. C. Seeing biological motion. *Nature* **395**, 894–896 (1998).
123. Logothetis, N. K., Pauls, J., Augath, M., Trinath, T. & Oeltermann, A. Neurophysiological investigation of the basis of the fMRI signal. *Nature* **412**, 150–157 (2001).
124. Grill-Spector, K. & Malach, R. fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychol.* **107**, 293–321 (2001).
125. Poggio, T. A theory of how the brain might work. *Proc. Cold Spring Harb. Symp. Quant. Biol.* **55**, 899–910 (1990).
126. Wallis, G. & Rolls, E. T. Invariant face and object recognition in the visual system. *Prog. Neurobiol.* **51**, 167 (1997).
127. Fox, R. & McDaniel, C. The perception of biological motion by human infants. *Science* **218**, 486–487 (1982).
128. Cavanagh, P., Labianca, A. P. & Thornton, I. M. Attention-based visual routines: sprites. *Cognition* **80**, 47–60 (2001).
- A psychophysical study using a visual search experiment to show that biological motion recognition in more complex tasks requires attention.**
129. Shiffrar, M. & Freyd, J. J. Apparent motion of the human body. *Psychol. Sci.* **1**, 257–264 (1993).
- A psychophysical investigation of the perceived paths of apparent motion between two photographs of human body configurations. For small ISIs, motion along the shortest paths is perceived, so the percept is dominated by low-level motion perception. For long ISIs, subjects perceive the biologically possible paths, indicating top-down influences from motion recognition.**
130. Umiltà, M. A. *et al.* I know what you are doing: a neurophysiological study. *Neuron* **31**, 155–165 (2001).
131. Verfaillie, K. & Daems, A. Representing and anticipating actions in vision. *Vis. Cogn.* **9**, 217–232 (2002).
132. Bassilli, J. N. Facial motion in the perception of faces and of emotional expression. *J. Exp. Psychol. Hum. Percept. Perform.* **4**, 373–379 (1978).
133. Poizner, H., Bellugi, U. & Lutes-Driscoll, V. Perception of American Sign Language in dynamic point-light displays. *J. Exp. Psychol. Hum. Percept. Perform.* **7**, 430–440 (1981).
134. Runeson, S. & Frykholm, G. Visual perception of lifted weight. *J. Exp. Psychol. Hum. Percept. Perform.* **7**, 733–740 (1981).
135. Verfaillie, K. Perceiving human locomotion: priming effects in direction discrimination. *Brain Cogn.* **44**, 192–213 (2000).
- A sound review of the experimental studies on biological motion perception that argues in favour of a view-based representation of biological motion patterns.**
136. Goddard, N. H. *The Perception of Articulated Motion: Recognizing Moving Light Displays*. Thesis, Dept. of Computer Science, Univ. of Rochester (1992).
- A connectionist model for biological motion recognition that proposes a similar architecture to the model used in this review.**
137. Little, J. J. & Boyd, J. Recognizing people by their gait: the shape of motion. *J. Videre* **1**, 1–32 (1998).
138. Rosenblum, M., Yacoob, Y. & Davis, L. S. Human expression recognition from motion using a radial basis function network architecture. *IEEE Trans. Neural Networks* **7**, 1121–1138 (1996).
139. Cedras, C. & Shah, M. Motion-based recognition: a survey. *Image Vis. Comp.* **13**, 129–155 (1995).
140. Gavrilá, D. M. The visual analysis of human movement: a survey. *Comp. Vis. Image Underst.* **73**, 82–98 (1999).
141. Aggarwal, J. K. & Cai, Q. Human motion analysis: a review. *Comp. Vis. Image Underst.* **73**, 428–440 (1999).
- References 139–141 review computer vision methods for the recognition of biological movements.**
142. Song, Y., Goncalves, L., di Bernardo, E. & Perona, P. Monocular perception of biological motion in Johansson displays. *Comp. Vis. Image Underst.* **81**, 303–327 (2001).

Acknowledgements
 We thank R. Blake, D. Perrett, T. Sejnowski, C. Koch, S. Edelman, I. Bülthoff, E. Curio, Z. Kourtzi, M. Riesenhuber, P. Sinha, I. Thornton and L. Vaina for comments; A. Benali, Z. Kourtzi and C. Curio for help with data acquisition; and the Max-Planck Institute for Biological Cybernetics, Tübingen, for providing support. M.G. was supported by the Deutsche Forschungsgemeinschaft, Honda R&D Americas Inc. and the Deutsche Volkswagen Stiftung. T.P. is supported in part by the Eugene McDermott chair. Research at CBCL was sponsored by the National Science Foundation. Additional support is provided by Central Research Institute of Electric Power Industry, Eastman Kodak Company, DaimlerChrysler AG, Compaq, Komatsu Ltd., NEC Fund, Nippon Telegraph & Telephone, Siemens Corporate Research Inc., Honda and Toyota.

 Online links

FURTHER INFORMATION
ARL homepage: [www.http://www.uni-tuebingen.de/uni/kn/ar/arl/home.html](http://www.uni-tuebingen.de/uni/kn/ar/arl/home.html)
Biological Motion Demonstrations: http://www.biols.susx.ac.uk/home/George_Mather/Motion/index.html
Randolph Blake: <http://www.psy.vanderbilt.edu/faculty/blake/BM/BioMot.html>
BioMotionLab: <http://www.bml.psy.ruhr-uni-bochum.de/index.htm>
CBCL homepage: <http://www.ai.mit.edu/projects/cbcl-web-homepage/web-homepage.html>
MIT Encyclopedia of Cognitive Sciences: http://cognet.mit.edu/MITECS/computational_vision/object_recognition_animal_studies/structure_from_visual_information_sources/vision_and_learning
 Access to this interactive links box is free online.