Learning Representations of Animated Motion Sequences - A Neural Model

Georg Layher (georg.layher@uni-ulm.de)
Institute of Neural Information Processing
Ulm University
James-Franck Ring, 89069 Ulm, Germany

Martin Giese (martin.giese@uni-tuebingen.de)
Section for Computational Sensomotorics
University Clinic Tübingen
Frondsbergstraße 23, 72074 Tübingen, Germany

Heiko Neumann (heiko.neumann@uni-ulm.de)
Institute of Neural Information Processing
Ulm University
James-Franck Ring, 89069 Ulm, Germany

Abstract

The detection and categorization of animate motions is a crucial task underlying social interaction and perceptual decision-making. Neural representations of perceived animate objects are built in the primate cortical region STS which is a region of convergent input from intermediate level form and motion representations. Populations of STS cells exist which are selectively responsive to specific animated motion sequences, such as walkers. It is still unclear how and to which extent form and motion information contribute to the generation of such representations and what kind of mechanisms are involved in the learning processes. The paper develops a cortical model architecture for the unsupervised learning of animated motion sequence representations. We demonstrate how the model automatically selects significant motion patterns as well as meaningful static form prototypes characterized by a high degree of articulation. Such key poses are selectively reinforced during learning through a cross-talk between the motion and form processing streams. Next, we show how sequence selective representations are learned in STS by fusing static form and motion input from the segregated bottom-up driving input streams. Cells in STS, in turn, feed their activities recurrently to their input sites along top-down signal pathways. We show how such learned feedback connections enable making predictions about future input as anticipation generated by sequence-selective STS cells. Network simulations demonstrate the computational capacity of the proposed model by reproducing several experimental findings from neurosciences and by accounting for recent behavioral data. **Keywords:** animated motion representation; implied motion; neural model; unsupervised learning; feedback.

Introduction

Animated movements in actions, like walking, turning, etc., can be robustly detected from video sequence input and predictions about future occurrences can be derived from such spatio-temporal patterns. Giese & Poggio (Giese & Poggio, 2003) proposed a hierarchical feedforward network architecture that aims at explaining the computational mechanisms underlying the perception of biological motion, mainly from impoverished stimuli such as point-light walkers. In this paper, we propose a new learning-based hierarchical model for analyzing animated motion sequences. Prototypes in the form and motion pathways are established using a modified Hebbian learning scheme. We suggest how snapshot prototypes are automatically selected from continuous input video streams utilizing features from the motion pathway which are indicative for the occurrence of specific snapshots with strongly articulated configurations, serving as key poses. Sequence-selective representations of articulated motions in cortical STS are driven jointly by input activations from both motion and form prototypes. In addition, feedback connections are learned to enable STS neurons predicting expected input from form selective IT and motion sensitive MST. We argue that for inputs presenting articulated postures without continuing motion, STS representations are fed by the corresponding snapshot prototype activations (Jellema & Perrett, 2003). In turn, STS will send feedback to stages in the segregated pathways for form as well as motion processing. Stationary images which depict articulated postures, consequently generate effects of implied motion, which have been shown in functional magnetic resonance imaging (fMRI) studies (Kourtzi & Kanwisher, 2000). We will argue here, that this can be accomplished by the proposed model through the action of fusing bottom-up input, driven by snapshot representation only, and the activated sequence representations sending feedback to *both* form and motion representations, thus amplifying motion representations even if no direct motion input is present.

Several computer vision approaches have been proposed for performing action recognition using different processing strategies of combining form and motion information. These approaches build upon the hierarchical architecture proposed by Poggio and coworkers which aims at defining a framework for form processing in the cortical ventral pathway (Riesenhuber & Poggio, 1999). Extensions of the form processing model to analyze motion information responses in a separate pathway, like the Giese-Poggio model, have been suggested in e.g. (Schindler & Van Gool, 2008). Here, the relative contributions of form and motion features to the classification of actions have been investigated. Details of the motion processing cascade alone have been studied in more detail in (Escobar & Kornprobst, 2012). Here the authors contributed further evidence that detecting motion contrasts in sequences of animated motion is useful to distinguish action classes. In all these proposed models, the mechanisms for hierarchical motion (and form) processing are predefined and learning only occurs at the level of a final classifier to distinguish given categories. It still remains unclear to a large extent, how the motion and form prototypes (e.g., in cortical areas MST and IT, respectively) and the sequence-selective pattern representations in STS interact and which features are

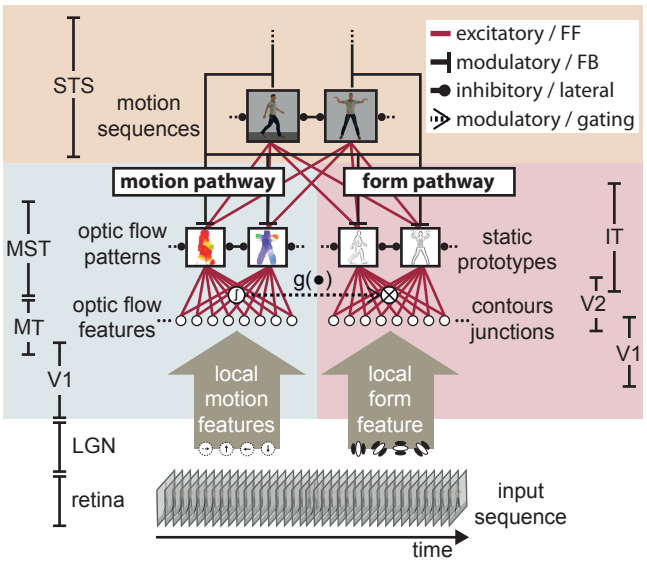

Figure 1: Overview of the model architecture. The model consists of two separate processing streams, the motion and the form pathway, both converging into model area STS. Static form prototypes in area IT, as well as optic flow patterns in area MST are learned using an unsupervised Hebbian mechanism. A motion driven reinforcement signal between the two pathways is used to steer the learning of the IT prototypes. After the suppression of cells with low activities, IT and MST cells propagate into area STS, where sequence-selective cells learn corresponding spatio-temporal activity patterns using a similar Hebbian learning rule. In addition, the sequence-selective cells learn the output weights back to the segregated form and motion prototypes, that stabilizes the input processing and activity fusion.

used for learning. How can feature representations be learned automatically from given input streams at different levels of the distributed action sequence representations? Also no top-down influences have been considered so far and how such connectivity patterns may transfer different information between pathways to generate proper predictions concerning future input configurations.

Model Architecture

The hierarchical model proposed here consists of two separate visual pathways for segregated form and motion processing as inspired by the work of (Giese & Poggio, 2003) and extends it by combining it with models for the hierarchical feedforward and feedback processing of motion and form along the dorsal and the ventral pathway (Bayerl & Neumann, 2004; Weidenbacher & Neumann, 2009). Intermediate level form representations (in model IT) and prototypical optical flow patterns (in model MST) are established using a modified competitive Hebbian learning scheme with convergent weight dynamics. The two separate hierarchical learning approaches are influenced partly by the work of Rolls and collaborators (Rolls & Milward, 2000), in which the authors have suggested that layered neuronal structures arranged in a hierarchy with increasingly larger connectivity kernels can learn invariant representations of objects and specific motion patterns. Here, we propose how such learning in a hierarchy can be utilized for learning sequence-selective representations of animated movement prototypes from convergent form and motion input. In addition, we suggest how a motion-driven reinforcement mechanism automatically selects relevant snapshots in the form path from video input streams. The activities of the prototypical form and motion cells converge in the model complex STS, where correlated temporal activations for specific sequences are learned. Sequence-selective representations are established by combined bottom-up and top-down learning, both based on modified Hebbian mechanisms. An overview of the model is shown in Fig. 1. The details are outlined below.

Form and Motion Processing

Processing the raw input data utilizes an initial stage of orientation and direction selective filtering (in model area V1). These responses are fed into separated pathways which are selective to static form representations (areas V2 and IT) and characteristic optical flow patterns (areas MT and MST). We use single compartment model neurons with gradual activation dynamics. The membrane potential of individual model neurons is calculated by conductance-based mechanisms of feed-forward integration of excitatory and inhibitory feeding input and a passive leakage. The potential can be enhanced by a gating mechanism to amplify the efficacy of the current potential by a matching top-down feedback signal. The membrane potential is finally regulated by a gain control mechanism that leads to activity normalization for a pool of neurons through mutual divisive inhibition. These mechanisms are summarized in a three-stage hierarchy of processing that includes input filtering, modulatory feedback, and pool normalization. The output of a cell is defined by a signal function which converts the membrane potential into a firing rate, or activity. Such model cells are grouped into layers which form abstract models of cortical areas.

Learning of Form and Motion Prototypes

First, we investigated how intermediate level feature representations can be learned in a biologically plausible fashion by exposing the network architecture with realistic input sequences. In order to generate feature representations of complex form and motion patterns we employ an unsupervised learning mechanism based on a modified Hebbian learning scheme. The modification stabilizes the learning such that the growing of weight efficacies is constrained to approach (bounded) activity levels of the input or the output activation. Motivated by the invariance properties observed by (Wallis & Rolls, 1997) we combined the modified Hebbian learning mechanism with a short-term memory trace of prolonged activity of the pre- or the post-synaptic cells (*trace rule*). The adaptation of weightings is controlled by post-synaptic cells

which, in turn, mutually compete for their ability to adjust their incoming connection weights. The particular details as well as the particular variations of the core architecture are explained below.

Hebbian learning in the form and motion pathways. In order to select the image regions that are fed to the learning of prototype representations a region of interest (ROI) is defined which represents a bounding box around the target object. Features within the target region are selected for learning feedforward connection weights in the form and the motion pathway, respectively. We employ the modified Hebbian learning rule

$$\Delta w_{ji}^{FF,s} = \eta_s \cdot \bar{v}_i^{post} \cdot (u_j^{pre} - \bar{v}_i^{post} \cdot w_{ji}^{FF,s}) \qquad (1)$$

where $\Delta w_{ji}^{FF,s}$ represents the discretized rate of change in the efficacy of the weighted connections with the learning rate η_s; $s \in \{form, motion\}$ indicates that the same core mechanisms are devoted to learning in the form and motion pathway, respectively. The variables $u_j^{pre} = f(x_j)$ and $v_i^{post} = f(y_i)$ are the firing rates driven by the membrane potential of pre- and post-synaptic cells, henceforth denoted as activity. The activity $\bar{v}_i$ of the post-synaptic cell is calculated by the temporal trace rule $\bar{v}_i^t = (1-\lambda)\bar{v}_i^{t-1} + \lambda v_i^t$, $0 < \lambda < 1$ (Földiák, 1991). The trace rule (see also (Wallis & Rolls, 1997; Rolls & Milward, 2000)) has been proposed to incorporate a short-term memory function for the cells to keep their activation over a short temporal window while adapting their weights. The term in brackets on the r.h.s. of learning equation 1 serves as a biologically plausible mechanism to bound the growth of the cells' input synaptic weights (Oja, 1982). The post-synaptic cells (with activity $\bar{v}_i^{post}$) which gate the learning of their respective input weights are arranged in a competitive layer of neurons competing for the best matching response and their subsequent ability to adapt their kernel of spatial input weights. In a nutshell, the layer of post-synaptic neurons competes to select a winning node for a given input presentation which, in turn, is allowed to automatically adapt their incoming (instar) synaptic weights. The temporal trace (or short-term memory) establishes that categories learn their average input over a short temporal interval thus allowing small pertubations for the changing input signals.

Reinforcing snapshot learning. The Giese-Poggio model (Giese & Poggio, 2003) suggests that sequence selectivity for biological motion recognition is driven by sequences of static snapshots. While the original model relies on snapshots that were regularly sampled temporally, we suggest a mechanism of how snapshots corresponding to strongly articulated poses can be selected automatically. Such snapshot representations are learned in the form channel by utilizing a gating reinforcement signal which is driven by the complementary representation of motion in the dorsal stage MT/MST. Formally, the weighted integration of motion energy over a given neighborhood is calculated by

$$m_e = \int_\Omega u_\phi(\mathbf{x}) \cdot \Lambda(\mathbf{x}) d\mathbf{x} d\phi \qquad (2)$$

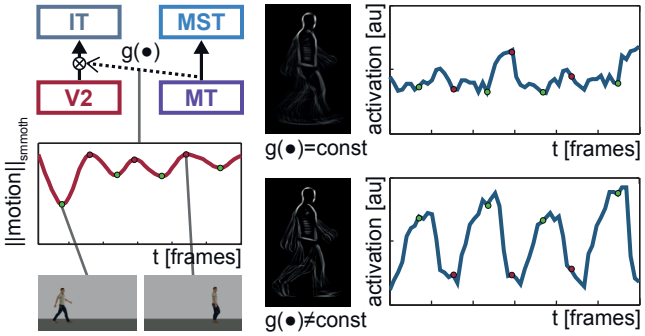

Figure 2: IT prototypes trained using disabled and enabled reinforcement signal. Minima and maxima in motion energy correspond to articulated and non-articulated postures (bottom left). Continuous learning of IT prototypes leads to activation profiles with low selectivity (top right). Motion driven reinforcement leads to IT prototypes which signal snapshot poses in synchrony with the gait (bottom right; for details, see text).

with $\Lambda(\bullet)$ denoting a spatial kernel for weighting the relative contribution of motion responses $u_\phi(\bullet)$ at spatial locations $\mathbf{x}$ in the 2-D image plane and with direction selectivity ϕ.[1] The motion energy signal itself is a function of time which is used to steer the instar learning in the form pathway. We suggest that different subpopulations of static form, or snapshot, representations can be learned that correspond to either weakly or strongly articulated postures. Here, we focus on snapshot poses corresponding to highly articulated postures with signatures of maximum limb spreading. Motion energy at limbs drops during phases of high articulation when their apparent direction of motion reverses. We incorporate the function $g(\bullet)$ to control a vigilance in snapshot learning to favor form inputs which co-occur with local motion energy minima, i.e. when $\partial_t m_e = 0$, given that $\partial_{tt} m_e > 0$. In the weight adaptation, $\Delta w_{ji}^{FF,form}$ in Eqn.1, the learning rate is now gated by the motion dependent reinforcement, $\eta_{form} \cdot g(m_e)$ which leads to the revised learning rule

$$\Delta w_{ji}^{FF,form} = \eta_{form} \cdot g(m_e) \cdot \bar{v}_i^{post} \cdot (u_j^{pre} - \bar{v}_i^{post} \cdot w_{ji}^{FF,form}). \qquad (3)$$

Learning of Sequence-Selective Representations

Categorial representations in the form and motion pathway, namely in IT and MST, which were learned at the previous stage, feed forward their activations to the stage of STS. In order to stabilize the representations and activity distributions, even in the case of partial loss of input signals, the STS sequence-selective representations send top-down signals to their respective input stages.

[1] For whole body motion considered here, we simply integrated the motion energy over the entire ROI without subdividing the image region. An analysis at smaller scales might necessitate an integration over smaller overlapping patches.

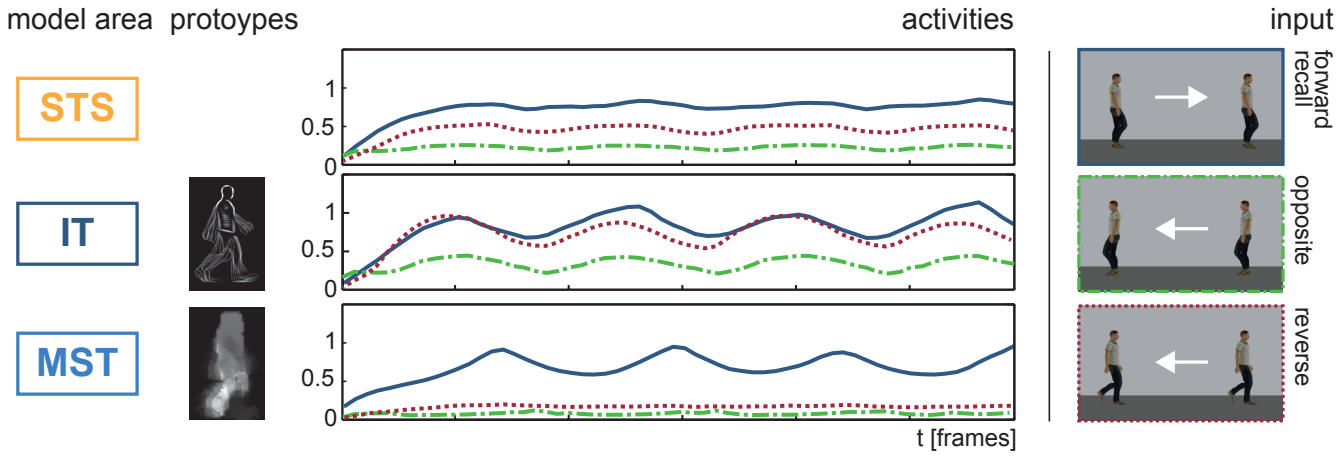

Figure 3: Response behavior of IT snapshot neurons, MST motion pattern neurons, and sequence-selective STS cells trained by video input for a walker moving from left to right. Activations in the model areas are shown for different input conditions for recall of the training sequence (top), opposite walker movement (middle), and walker displayed in reverse motion (bottom). Line styles and colors encode the input test cases on the right. For details and brief discussion, see text.

Learning of feedforward connections. Prototypical representations with spatio-temporal sequence selectivity are suggested to exist in the cortical STS complex where both form and motion pathways converge. The selectivities of model STS neurons are learned by again using a modified Hebbian instar learning mechanism similar to the separate learning of form and motion prototypes (Eqn.1),

$$\Delta w_{ji}^{in,FF} = \eta_{seqFF} \cdot \bar{v}_i^{post} \cdot (u_j^{pre} - \bar{v}_i^{post} \cdot w_{ji}^{in,FF}). \qquad (4)$$

The weighting kernel $w_{ji}^{in,FF}$ represents convergent IT $\rightarrow$ STS and MST $\rightarrow$ STS bottom-up input to a post-synaptic STS cell (instar). η_{seqFF} denotes the learning rate and u_j and v_i are the firing rates of the pre- and post-synaptic neurons, respectively (the post-synaptic activity is again calculated via a temporal trace mechanism). The pre-synaptic activity for the receiving model STS cells are generated by concatenating form and motion output activations, namely $\mathbf{u} = \mathbf{u}^{IT} \cup \mathbf{u}^{MST}$.

Learning feedback connections. An important component is that sequence-selective prototypes in STS in turn learn the output weights back to the segregated form and motion prototype representations, namely STS $\rightarrow$ IT + MST. Unlike the FF learning mechanisms, the learning here is gated by the pre-synaptic cell (in STS) for their top-down weights, which reads

$$\Delta w_{ji}^{out,FB} = \eta_{seqFB} \cdot \bar{v}_i^{pre} \cdot (u_j^{post} - w_{ji}^{out,FB}) \qquad (5)$$

with the same components as in the bottom-up learning formalism in Eqn.4. Bottom-up and top-down learning schemes slightly differ in the definition of the competitive terms (in brackets). In the feedback learning we employ a difference term between post-synaptic activity and the weighting, $u_j^{post} - w_{ji}^{out,FB}$, omitting the additional weighting of the connectivity strength via the pre-synaptic activity as in the Oja rule. In steady-state each of the connection strengths emanating from STS cells assumes a value corresponding to the post-synaptic activity distribution, which defines the current input activation. Given an STS cell with attraction $\bar{v}_i^{pre}$, the top-down weight vector approaches $\mathbf{u}^{post} = \mathbf{w}_i^{out,FB}$, thus learning the expected average input. Combined with the temporal trace, this establishes a representation in which each STS sequence-selective prototype encodes and memorizes in its weight pattern the expected driving input activity pattern configuration from the form and motion pathway. Such a top-down weighting pattern can then be used to generate predictions concerning the expected future input given the current maximally activated prototype at the STS level.

Results

The model has been tested in various computational experiments, not all of which we can present here. In a first experiment, we probed the properties of snapshot selection from the input streams and their signature concerning static articulations. The latter property has been motivated by the fact that extremal articulation indicates configurations of implied motion, in turn, predictive for future motions. Results shown in Fig. 2 demonstrate that input activations (in V2) with strongly articulated shapes cohere with local motion minima. Such minima drive the reinforcement signal for learning whole body form prototypes. Temporal response signatures for IT prototypes are shown for disabled reinforcement ($g(m_e) = 1$, and when it is enabled ($g(m_e)$ monotonically decreasing function of m_e).

We studied the response properties of STS representations and their motion sequence selectivity. There, a prototypical sequence-selective representation is learned for a walker that is traversing from left to right. After training of form, motion and sequence representations, the network is probed by

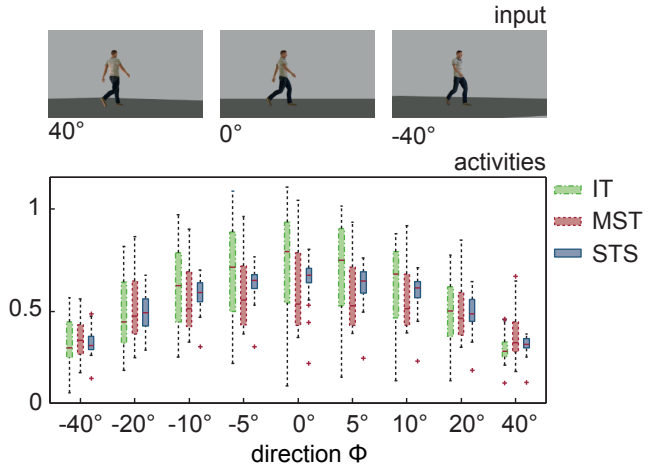

Figure 4: Response tunings of models cells in area IT (snapshots), MST (motion patterns), and STS (sequence-selective patterns) after training. Category representations have been learned for a walker moving along horizontal direction for $\phi = 0°$. Activities of prototypical cells are shown (bottom) which were probed by different inputs with varying movement directions, i.e. walkers approaching or receding at different angles with respect to the horizontal reference axis (top). Data has been summarized into box plots showing the response variabilities of models cells as well as the monotonic decline in response for deviations from the target tuning. The tuning width at half maximum response is around $\pm 40°$. The variance of the MST / IT prototypes decreases towards larger deviations, depicting the loss of response selectivity of prototypes to different parts of a walkers gait.

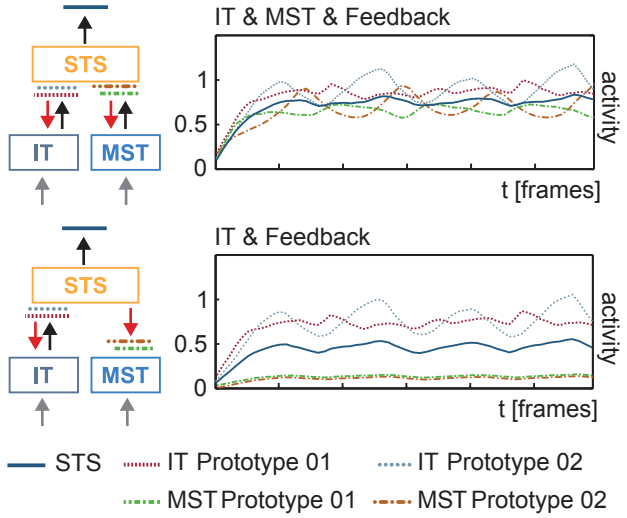

Figure 5: Selective removal of interconnections (lesioning). The model was trained using the same walking sequence as in the second experiment (see Fig. 3 / *forward recall*). The model was left untouched to provide a reference (top). Bottom-up (feedforward) connections between area MST and STS were removed, preventing any motion-related signal being propagated to STS (bottom). The amplitude of the IT prototype activities remains almost the same, whereas the sequence-selective STS cell responds only at about half-magnitude (because of the missing support from the motion pathway). Note the feedback activities propagated from STS to MST optical flow pattern prototypes. We argue that this reflects the induction of increased fMRI BOLD response in human MT+ following the presentation of static implied motion stimuli.

three different movement scenarios: a forward moving walker with same profile and movement direction as in the training phase (*recall*), a forward moving walker traversing from right to left (*opposite*), and a backward moving walker (*reverse*). Form/motion prototypes and the sequence representation are triggered maximally in the *recall* case while in the *opposite* case form and motion prototypes only respond minimally, and so do the sequence-selective cells. In the *reverse* case the form prototypes selectively match the input at high articulation configurations, while the motion responses remain minimal. As a consequence, the sequence-selective representations respond at an intermediate level (Fig. 3). This evidence is in line with the experimental findings by (Oram & Perrett, 1996) and recent observations by (Singer & Sheinberg, 2010).

We further investigated the direction tuning of the sequence-selective prototypes. Here, we configured different walkers with varying movement directions and speeds with reference to a previously learned representation of a rightward moving walker at a speed of 1 m/s. Walking directions in the test cases were rotated by $\pm\{5°, 10°, 20°, 40°\}$. Model simulations result in a direction tuning of STS cells with half amplitude of approximately ± 40 deg (Fig. 4). IT and MST cells, on the other hand, also show a drop in response but have a much larger variability.

In an additional experiment we selectively lesioned of the model architecture, particularly investigating the effects of extinguishing connections between model areas and the activity flow between learned representations (Fig.5). The fully connected model with learned IT / MST and STS feedforward and feedback connections was used as reference. When bottom-up connections from motion input (MST) were cut off the sequence-selective neuron responses in STS drop to approximately half their response amplitude. Feedback from STS invokes an amplification of activities in IT and MST representations. We observe that FF activation from IT alone can drive sequence neurons. Snapshot representations in IT drive the STS sequence neurons which, in turn, send feedback signals to the stages of IT and MST prototype representations. In the motion pathway such feedback elicits an increase in presynaptic activation. We argue that this reflects the induction of increased fMRI BOLD response in human MT+ following the presentation of static implied motion stimuli (Kourtzi & Kanwisher, 2000).

Discussion and Conclusion

We propose a biologically plausible model for the learning of animated motion sequences. The model builds upon neurophysiological evidence about the cortical sites and specific neuronal representations which contribute to articulated motion and implied motion perception. The main contributions of the paper are several-fold: First, we suggest how prototype representations in the form and motion pathways, namely in model cortical areas IT and MST, can be established on the basis of probing the model architecture by sequences containing animated motions. Learning mechanisms are based on modified Hebbian schemes which are stabilized through a trace mechanisms and the incorportion of an objective function taking the weight kernel saturation into account. Second, we suggest that sequence-selective cells in model area STS are learned by using the same learning mechanisms but now by combining the responses of intermediate level representations in the form and motion pathways. Third, the learning of articulated poses (snapshots) is controlled by a reinforcement mechanism that enables Hebbian learning in the form pathway through cross-pathway motion-form interaction. Given an animated motion sequence, snapshots are automatically selected as key poses corresponding to strong body pose articulations. Finally, the sequence-selective cells in model STS project to their respective input representations in the form and motion pathways. These feedback connections are again learned by a Hebbian mechanism. Together, the feedforward and the feedback interactions establish a loop of recurrent processing to stabilize the patterns of form, motion, and sequence representation. Via feedback, model STS cells generate a predictive signal through the backward connections' weights to encode the expected matching input that is suitable to match the currently activated sequence pattern. Together with the newly proposed feedback mechanism the model is able to account for various experimental findings, in particular, the ability to infer and predict future motion sequence development from articulated postures (implied motion). Importantly, cells in STS are responsive to *both* motion as well as static form (Oram & Perrett, 1996). The model predicts that the presentation of static key poses from previously learned sequences alone leads to enhanced activation in STS sequence selective neurons as observed in (Jellema & Perrett, 2003). The model also hypothesizes how the presentation of static articulated poses leads to the emergence of predictive motion perception and enhanced neural activations in the motion pathway (Kourtzi & Kanwisher, 2000). Furthermore, learned sequence-selective prototype representations have direction tunings in response to walkers in the range of ± 40, similar to those reported in (Perrett et al., 1989)). Once again, the model makes a testable prediction that articulated poses represent the snapshot frames that have been suggested by (Giese & Poggio, 2003) and that have recently been tested experimentally by (Singer & Sheinberg, 2010).

Acknowledgements

GL and HN have been supported by the SFB Transregio 62 funded by the German Research Foundation (DFG).

References

Bayerl, P., & Neumann, H. (2004). Disambiguating visual motion through contextual feedback modulation. *Neural Comput*, *16*(10), 2041–2066.

Escobar, M., & Kornprobst, P. (2012). Action recognition via bio-inspired features: The richness of center–surround interaction. *Comput Vis Image Und*, *116*(5), 593–605.

Földiák, P. (1991). Learning invariance from transformation sequences. *Neural Comput*, *3*(2), 194–200.

Giese, M., & Poggio, T. (2003). Neural mechanisms for the recognition of biological movements. *Nat Rev Neurosci*, *4*(3), 179–192.

Jellema, T., & Perrett, D. (2003). Cells in monkey STS responsive to articulated body motions and consequent static posture: a case of implied motion? *Neuropsychologia*, *41*(13), 1728–1737.

Kourtzi, Z., & Kanwisher, N. (2000). Activation in human MT/MST by static images with implied motion. *J Cognitive Neurosci*, *12*(1), 48–55.

Oja, E. (1982). Simplified neuron model as a principal component analyzer. *J Math Biol*, *15*(3), 267–273.

Oram, M., & Perrett, D. (1996). Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey. *J Neurophysiol*, *76*(1), 109–129.

Perrett, D., Harries, M., Bevan, R., Thomas, S., Benson, P., Mistlin, A., ... Ortega, J. (1989). Frameworks of analysis for the neural representation of animate objects and actions. *J Exp Biol*, *146*(1), 87–113.

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nat Neurosci*, *2*, 1019–1025.

Rolls, E. T., & Milward, T. T. (2000, November). A model of invariant object recognition in the visual system: Learning rules, activation functions, lateral inhibition, and information-based performance measures. *Neural Comput*, *12*(11), 2547–2572.

Schindler, K., & Van Gool, L. (2008). Action snippets: How many frames does human action recognition require? In *CVPR 2008* (pp. 1–8).

Singer, J., & Sheinberg, D. (2010). Temporal cortex neurons encode articulated actions as slow sequences of integrated poses. *J Neurosci*, *30*(8), 3133–3145.

Wallis, G., & Rolls, E. (1997). Invariant face and object recognition in the visual system. *Prog Neurobiol*, *51*(2), 167–194.

Weidenbacher, U., & Neumann, H. (2009). Extraction of surface-related features in a recurrent model of V1-V2 interactions. *PloS ONE*, *4*(6), e5909.