



HAL
open science

Interactions of motion and form in visual cortex: a neural model

Cornelia Beck, Heiko Neumann

► **To cite this version:**

Cornelia Beck, Heiko Neumann. Interactions of motion and form in visual cortex: a neural model. Deuxième conférence française de Neurosciences Computationnelles, "Neurocomp08", Oct 2008, Marseille, France. hal-00331597

HAL Id: hal-00331597

<https://hal.science/hal-00331597>

Submitted on 17 Oct 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INTERACTIONS OF MOTION AND FORM IN VISUAL CORTEX – A NEURAL MODEL

Cornelia Beck and Heiko Neumann
Institute for Neural Information Processing, University of Ulm
Oberer Eselsberg, 89069 Ulm, Germany
cornelia.beck@uni-ulm.de, heiko.neumann@uni-ulm.de

ABSTRACT

In this work we present a neural model simulating parts of the motion and the form pathway of the visual cortex. It is shown how the visual features motion, disparity, and form that are represented in a distributed way in areas V1, V2, and MT mutually interact at several levels to share information without the need of explicit place-like coding. In particular, we address the issue of 2D extrinsic motion cues generated at occlusions that have to be treated differently than 2D intrinsic motion features of the same object. We suggest that here information of the form channel, namely the indication of a junction, is necessary to achieve a correct percept in the motion pathway. Furthermore, we investigated the question of how a percept of either pattern or component motion is generated in a scenario of moving bars that only differ in the presence or absence of occlusions, like in the chopstick display. Therefore, we propose different roles for various kinds of MT cells that are involved in the interactions with the form pathway, simulating purely integrative cells tuned to motion and to motion and stereo, but also contrast cells responding strongly to opponent motion in the surround. The model simulations reproduce psychophysical and neurophysiological results of the chopstick as well as of the barberpole illusion. The temporal course of the dominant motion percept generated by the iterative interplay between motion and form pathway is in line with data of ocular following responses in primates and humans.

KEY WORDS

Neural model, vision, motion, form, feedback, chopstick illusion

1. Introduction

Surfaces moving in space are characterized by feature conjunctions of lightness, velocity, relative depth, etc. Related feature detection processes activate certain representations in the brain whose structure as well as the possible interaction of the underlying neural processes is a topic of intense investigation. Evidence suggests that early and mid-level processing in area V1, V2, and MT of static and dynamic visual information is organized along parallel streams with mutual interactions at different levels [1]. One advantage of this architecture is that features can be coded in a basically distributed fashion while mutual connections between processing stages allow the exchange of information. In MT for example, cells tuned to both motion and

disparity exist [2]. Ponce et al. [3] showed that when areas in the form pathway (V2/V3) are deactivated, the disparity tuning of these MT cells is largely reduced, which is a clear hint to the interactions between the two pathways. The corresponding roles that different cell types found in MT play in this context is not yet clear. There are several problems to solve, like the segmentation of the objects of a scene while at the same time integrating regions that belong together. The representation in lower areas should remain spatially localized despite feedback of higher areas with larger RF fields. Furthermore, ambiguous motion cues that appear along a 1D structure have to be resolved (“aperture problem”). At 2D structures, where unambiguous motion estimates can be computed, another difficulty appears: If the structure is intrinsic, i.e. belonging to the object, the motion estimated there is reliable and helps to solve the aperture problem. However, if the termination is extrinsic and thus belonging to a different object surface because of an occlusion, the feature response needs to be suppressed, as measured by [4]. We suggest that these issues can be solved by appropriate interactions between the two pathways, as presented in the following chapters. Also, the contribution of different kinds of MT cells within these mechanisms is addressed. We suggest how different percepts of global motion are generated for stimulus configurations in changing contexts, like in the chopstick experiment [5]. In model simulations we show how feedback from MT to V1 cell populations and mutual form-motion interactions between V2 and MT influence the generation of global percepts.

The question of how motion estimates for stimuli formed by several components is computed is also connected to the generation of ocular following responses (OFR). Evidence exists that they are directly influenced by cortical motion processing in areas MT/MST [6, 7]. Our model presented here represents a possible basis for motion segmentation to explain the time course of human and primate OFR for moving stimuli perceived as pattern motion like in the barberpole experiment [8, 9, 10].

2. Model

Our model consists of areas V1, V2, and MT to present the first stages along the dorsal and the ventral pathway as depicted in Fig. 1. We build upon the components of a previously developed model of motion processing using feedforward and feedback interaction between areas V1 and MT [11]. This model suggests a possible

explanation how the motion aperture problem is solved [12]. In the extended model proposed here, MT contains two kinds of integration cells, MT Motion cells that are only tuned to motion and MT Motion+Stereo cells tuned to both motion and disparity [2]. In addition, MT Motion contrast cells are included utilizing a center-surround mechanism with opposite direction tuning. These cells respond strongly if the velocities in the surround differ from the movement that is found in the center. The mainly integrative fashion of the forward processing in the motion path is indicated by increasingly larger receptive field sizes, with a ratio of approximately 1:5 for V1:MT Motion and MT Motion+Stereo and 1:1.25 for MT Motion : MT Motion contrast.

The stages along the form pathway, namely model areas V1 and V2, are modelled on the basis of a former model of long-range grouping and texture boundary formation [13]. Both form and disparity tuned cells are simulated. V1 Form cells detect local luminance gradients of eight different orientations, V1 Stereo cells are activated by shifts in horizontal direction between the left and the right input image based on the responses of a correlation-based disparity detector (similar to the Reichardt detector used here for motion detection) that is tuned to different horizontal disparities. V2 Form and Stereo cells are then integrating the corresponding V1 inputs using bipole filters. Multiplicative combinations of V1 and V2 Form cells allow the detection of end-stop positions and possible X- or T-junctions (see Fig. 1 right, ratio of RF size V1:End-stop and V1:Bipole approx. 1:4). Model areas are modular in the sense that they are all built by a cascade of processing stages, including a filtering stage for driving input, a feedback signal path to modulate the input, and a stage of center-surround shunting competition. The dynamics of the individual stages is defined formally by first-order ordinary differential equations, utilizing single-compartment neuron models at the individual processing stages. In particular, we have

$$\partial_t v^{(1)} = -v^{(1)} + s^{FF} * \Lambda_{\sigma_1}^{(\mathbf{x}, space)} * \Psi_{\sigma_2}^{(\phi, velocity)} \quad (1)$$

$$\partial_t v^{(2)} = -v^{(2)} + (v^{(1)})^2 \cdot (1 + C \cdot z^{FB}) \quad (2)$$

$$\partial_t v^{(3)} = -Av^{(3)} + v^{(2)} - (E + v^{(3)}) \cdot \sum_{\phi} v^{(2)}. \quad (3)$$

The terms $v^{(1)}$, $v^{(2)}$, and, $v^{(3)}$ denote the activity within the three stages of the model area, the term s^{FF} in (1) denotes the driving input signal, while z^{FB} in (2) is the modulatory feedback signal; the functions Λ and Ψ in (1) denote weighting kernels in the spatial and the velocity domain, respectively, $*$ denotes the convolution operator, \mathbf{x} the spatial position, ϕ denotes velocities, and the constants A , C , and E in (2) and (3) adjust proper activity decay and strengths of feedback and lateral inhibition. In the results presented here, the steady-state-solutions of Eq. (1)-(3) are used to compute the neural activity. For this reason, we do not take into account the delay of neural responses between the different model areas.

We suggest that the mutual cross-pathway interactions between selective representations of motion and form/disparity, namely bidirectional V1/V2-MT and MT-V1/V2 processing, are based on modulatory interactions. The bidirectional mutual interactions between V1/V2 Stereo and MT Motion+Stereo, the feedback from MT Motion contrast cells, and the input from end-stop activity to MT Motion using a gain enhancement such as denoted in Eq. (2) all follow the linking principle proposed by Eckhorn et al. [14]. Further interactions between form and motion pathway are the input at possible X- and T-junctions. A tonic input that is not velocity selective is added to MT Motion cells after the feedback step – the net effect is a suppression at these positions due to the subsequent normalization process (see [15]). The influence of this interaction is still modulatory in a way that the activity is only reduced, but not completely inhibited. V2 Form information also influences the integration process in MT Motion. This means, that the feedforward activity of V1 is not simply added up in an isotropic way, but it is weighted according to the strength of the form activity in V2.

The model was tested with input images of size 120x90 pixels. For motion cells, we represented neurons responding to a spatial shift from -6 up to +6 pixels in both x and y direction, stereo cells were only separated in two layers for near and far stimuli, 8 different orientations were used in the form channel.

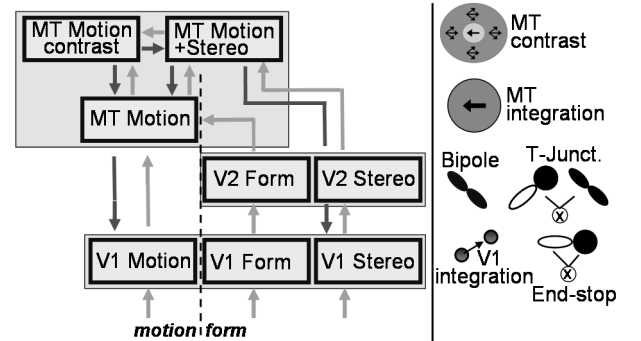


Fig. 1. Left: Neural model. Motion and form pathway are modelled simulating area V1, V2, and MT. Feedforward interactions are indicated using light gray arrows, modulatory interactions using dark gray arrows. The input to V1 subareas is the correlation response of an elaborated Reichardt detector in V1 Motion and V1 Stereo and oriented luminance contrasts for V1 Form. Right: Cell types. V1 cells simply integrate over a small area, V2 RFs have the shape of a bipole. End-stop positions and T-junctions are indicated by the multiplicative combination of the depicted subfields. In MT, both Motion and Motion+Stereo cells are integrating their input (influence of V2 Form on MT Motion not depicted here), whereas the Motion contrast cells have a different velocity tuning in the center and the surround.

Altogether, in different model areas according to their resolution and the tunings of the neurons, large numbers of neurons would have to be simulated. We used the principle of rank-/latency coding to achieve a sparse representation and an efficient computation of the neural activity [16]. To reduce the number of active neurons after every processing stage only the neurons

with a latency smaller than a certain threshold were kept for the simulations (only approx. 1% of the neural activity had to be represented in each layer). The latency of the neurons was computed using the model of quadratic integrate and fire neurons of Izhikevich [17]. Our results show that such a representation is sufficient to get simulation results corresponding to psychophysical and neurophysiological experiments.

3. Results

The complementary interaction of the processing streams in the model were initially tested using input sequences with two bars aligned like an “X” moving horizontally in opposite directions. This configuration is interesting because when two horizontal occluders are added at top and bottom of the bar endings the stimulus leads to a different percept switching from component motion to pattern motion. The bars are no longer perceived as two independently moving objects, but as one pattern moving in a coherent way. This effect, called the “chopstick illusion” [5], can be explained by the mutual interactions between the motion and the form pathway. In the first experiment we show the result for the chopstick illusion without occluders, where two bars are moving in the same depth plane in opposite directions.

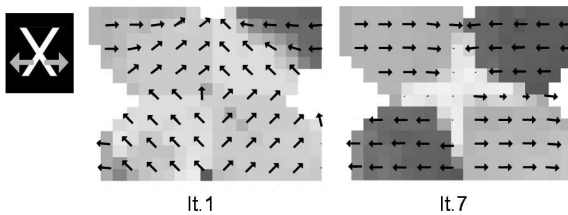


Fig. 2. Experiment 1: Mean velocity of area MT motion. Left: After the first iteration, only the line endings indicate the correct motion, along the edges normal flow is estimated indicated by the arrows orthogonal to the line contrasts. At the crossing upward motion is estimated due to the local movement of the junction formed by the two bars. Right: After 7 iterations MT Motion cells clearly achieve a segmentation and indicate the two correct independent motion directions of both bars. The two stimuli are segregated in two objects, one moving to the left, the other one to the right while the pattern motion at the crossing is suppressed. Note that no depth information is included in the input images. Therefore, at the crossing all the neurons tuned to one of the two component movements respond. As the sum of the left and right movement add up to a mean velocity of approx. 0, these positions are indicated in white/light gray (center right).

Initially, the motion estimates in area MT shown in Fig. 2 represent the true bar movement only at the bar endings. The 2D intrinsic terminators there achieve a correct motion estimate. In contrast, along the edges at the beginning normal flow is indicated, in the center the locally measured pattern motion is represented, the upward motion of the crossing itself. But during several iterations, the strong activity at the bar endings can propagate along the bars (Fig. 2 right). This effect is in particular facilitated by two mechanisms: First, the contrast cells respond strongly at the bar endings. In the center of the RF high input activity is present due to the unambiguous activity at the bar ending. In the large surround, the opponent motion of the other bar further

strengthens the activity. The strong activity of the contrast cells is fed back to MT integration cells and strengthening the component motion of each bar. Second, the activity in the central part indicating motion in upward direction is weakened due to input from V2 grouping cells indicating that at this position an X-junction appears.

We now add static occluders to the input images (experiment 2), to investigate the influence of the form channel when occlusions appear in the image. The human perception for this scenario is that the whole pattern formed by the two bars is moving in a coherent way. This phenomenon is replicated by our model results: As represented in Fig. 3 the overall motion after 7 iterations (right image) shows a uniform upward movement. Which mechanisms contribute to this effect leading to a completely different percept compared to experiment 1? After the first iteration (Fig. 3, left) the motion at the bar endings is already less pronounced than in experiment 1. The T-junction formed by the bar endings and the occluders leads to strong activity in the form channel that has an additive input in MT Motion for all velocities at this position. After the normalization in MT Motion, this results in reduced activity for the component motion. At the same time, MT contrast cells receive less input from the bar endings and, hereby, the feedback supporting the component motion is only small. These changes allow the pattern motion to propagate along the edges and hence to dominate the whole percept.

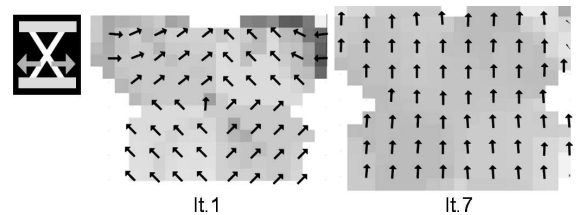


Fig. 3. Experiment 2: Mean velocity of MT Motion cells after 1 (left) and 7 (right) iterations for the chopstick scenario with occluders. At the beginning, the response of the MT Motion neurons is very similar to the case without occluders (cmp. Fig. 2), but there is less activity at the bar endings. After several iterations the pattern movement that both bars form together dominates the percept of the whole scenario.

The suggested model utilizes independent pathways and complex interactions, both within and across the pathways. In order to investigate the individual contributions of the components and their interactions in more detail we can selectively lesion the network model. Such lesions involve (a) the silencing of particular areas or sub-systems (by eliminating their computations and the resulting representations) and (b) the selective cutting of connections or reducing their impact in the simulations. Due to size limitations we cannot present a systematic evaluation of all individual contributions. We focus here on systematic impairments of computations by cutting modulatory feedback connections in model simulations and compare the results with those achieved by the intact model. For example, in Fig. 4 the results for experiment 1 and 2 are shown with feedback from MT Contrast cells having been eliminated. These cells respond strongly if

opposite motion directions are found in area MT in their RF center and surround. For the chopstick experiment with occluders the effect is negligible (Fig. 4 right), again global movement for the whole stimulus is indicated. However, without occluders the results look different: When the additional strengthening from top for the two bars moving in different directions is missing, the motion segregation is not successful. In the center, the pattern motion is still indicated, in the other regions the true motion direction is only partly achieved. This is an indicator for the functional role of the contrast cells in the model. We also tested the influence of V2 Form cells (results not shown). Here, the modulatory connections have been weakened by a factor of 0.2. Results look similar to those without MT contrast cell feedback depicted in Fig. 4. The model is no longer able to segregate two bars as the center motion is not suppressed sufficiently.

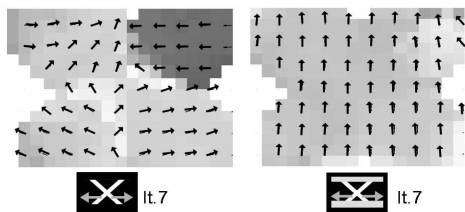


Fig. 4. Experiment 1 (left) and 2 (right) with feedback connections from MT contrast cells having been eliminated. The mean velocity estimated in MT is shown for the two pattern configurations after 7 iterations. For experiment 1, in contrast to the results with the full model connections, no clear segmentation of the two independently moving bars is achieved. The result for experiment 2 without this feedback is very similar to the former results (cmp. Fig. 3 right) indicating global motion for the whole stimulus.

In the third experiment we want to demonstrate the effects of the complementary interactions between disparity and motion coding. When the two bars are presented in two different depth planes, the bar in front should determine the motion in the center. At the same time, the bar in front should be presented as a continuous object in the V2 Stereo cells without an interception at the center where the two bars cross. The simulations showed that after only few iterations the representation of the near bar can be completed in V2 Stereo cells, whereas the second bar in the background is not completed. At the same time, the motion of the bar in front in MT Motion+Stereo is propagated into the central part as depicted in Fig. 5: In the first iteration mainly the normal flow activity prevails along the edges, after 7 iterations a clear segregation of the two independently moving objects is achieved with highest neural activity for the corresponding component direction. The movement of the bar in front has been propagated to the central part over few iterations. In addition, we probed the model in a fourth experiment with the barberpole stimuli [18], a scenario where diagonal grating moves in normal flow direction behind an invisible rectangular aperture leading to the illusory percept of movement in direction of the longer edge of the aperture. Also in this scenario, the interplay of local and global mechanisms of both

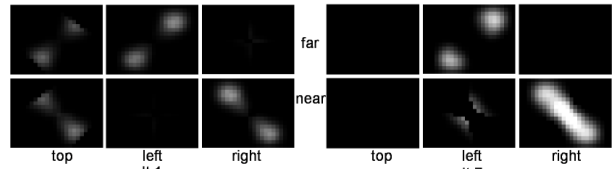


Fig. 5: Experiment 3: Activity of MT Motion+Stereo neurons tuned to upward, left and right motion for the far (top row) and the near (bottom row) plane. Left: After the first iteration, both the near and the far bar are not continuously represented, also many neurons tuned to the pattern motion respond along the edges (aperture problem, activity in other directions not shown). Right: After some iterations, the bar representation of the near bar is correctly completed: Neurons tuned to the near plane and movement to the right reach high activity all along the outline of the near bar. In contrast, the bar in the background is correctly not completed.

motion and form is crucial for the percept. In contrast to the chopstick illusion, no segmentation of single objects is necessary, but the integration of the overall response to the grating seems to determine the human percept. Our results show that the model response is coherent with the perceptual phenomenon measured in humans and macaques [18, 19]: (1) For a square aperture, the main activity in MT corresponds to the normal flow direction. (2) For a rectangular aperture that is elongated markedly, the motion measured at the intrinsic 2D terminators located along the longer axis influence the percept: A decision based on the MT activity leads to a motion direction between the normal flow direction and the flow direction of these terminators. (3) For a square aperture and occluders in horizontal direction, the response is dominated by the response of the vertical direction (results not shown).

4. Discussion

New contributions

We have developed a model for the early and intermediate stages of motion and form processing in visual cortex. It aims at revealing the function of cortical areas with neurons of different computational competences and their interactions to generate coherent task related representations of the input stimuli. The model is used to replicate perceptual results of psychophysical and neurophysiological experiments.

The new contributions of this work are in particular (i) the development of a model for interactions between the form and motion channel including form, motion and stereo features that attributes different roles for different kinds of MT cells, (ii) an explanation of different perceptual phenomena in one single model that has generic mechanisms in all its areas, (iii) the specification of a new way of integrating motion estimates in MT guided by form information, and (iv) a suggestion of how the interaction between the different areas allows complementary coding keeping a partially distributed representation of the features.

Related work

Other work addressing the issue of complementary coding of motion and form is the model of Berzhanskaya et al. [20], building on previous developments in [21], that emphasizes in particular the

detailed mechanisms in the processing areas up to the different layers within V1, MT, etc., and uses very large-spanning RF sizes in the higher levels like MT and MST. In contrast, we stress the combination of coarser, but still localized resolution context with a high-resolution form representation via feedforward and modulatory feedback. This means, that despite larger RF sizes in area MT, we keep a localized response in the lower areas. For example, in experiment 3 the MT Motion+Stereo cells contribute to the continuous representation of the near bar in V2 Stereo, where the high spatial resolution is preserved.

Another model for form-motion interaction was proposed by Lidén & Pack [22] where T-junctions explicitly inhibit motion signals. In our contribution, an excitatory additive signal of V2 form enhances all motion sensitive cells at locations of potential occlusions. This tonic input in velocity space leads to decreases of salient responses at this position: First, the activity of all motion sensitive cells is increased by the input, but the following normalization stage, in turn, shunts down the entire activity pattern. This mechanism provides the possibility for a highly activated neuron to keep its saliency even within the normalization process. It is hence consistent with the idea of only modulatory interactions between the two pathways.

Tlapale et al. [23] proposed an extension of the motion model by Bayerl & Neumann [11] to simulate the chopstick as well as the barberpole illusion. The major difference to the original motion model is the integration in MT Motion cells that is depending on the local luminance similarity and does not simply use spatial integration weighted with a gaussian function. Form and disparity information are not used. In our model we picked up the idea of an integration in MT Motion that is not purely isotropic. In contrast to their approach, we use the response of bipole filters in V2 Form to steer the integration process in MT Motion. This has the advantage that the integration is not based on simple, and often erroneous luminance differences, but on more reliable cues of the form pathway. This kind of integration is, e.g., an advantage in the case of the chopstick experiment with occluders: Due to the different orientations of the occluders and the bar endings, the integration will basically stop at the junction of the different objects. Hence, the motion integration for the different objects is more separated.

Structure and function of motion-form interaction

The model mechanisms presented in this work incorporate reentrant processing in the motion and form pathway, suggest various cross-pathway interactions, and discuss their consequences for building up a distributed representation of segregated or coherently moving forms. To show the functional contribution of the different areas and connections, we chose the chopstick and the barberpole experiment. The chopstick illusion represents a challenging task for correct segregation as depending on the context (occluders/no occluders) the movement of the whole stimulus is interpreted in different ways.

Without occluders the motion of the line endings indicating the correct component motion has to be

propagated and it has to suppress the central pattern motion. We suggest that the segregation of localized objects is supported by MT contrast cells that respond particularly strong for movements surrounded by opponent movement. Feedback from this area enhances MT Motion and Motion+Stereo cells which further feed back to V1 Motion cells. This mechanism supports the segmentation of different objects in the entire motion processing loop. The importance of the contrast cells gets visible, if we cut the feedback connection to V1/MT Motion cells (Fig. 4). Without their feedback, the two bars moving separately can no longer be segmented correctly. Support for their opponent movement is necessary to propagate the true movement from the endings all along the bars, also into the central part that has initially strong motion feature tracking signals in the vertical direction. In experiment 3 the same scenario was used, but the bars were represented in different depth planes. The results show that MT Motion+Stereo cells correctly achieve a segregation of the two bars due to the V2 Stereo information. Also here, the correct propagation along the bars depends on the MT contrast cells. The additional depth information allows the correct assignment of the central region: After some iterations it only represents the movement of the bar in front. Without depth information in this region the motion of both bars is indicated.

In the scenario with occluders, the disambiguation of the motion is driven by the central junction of the two bars with pattern motion, along the occluders the activities are suppressed. The input of the contrast cells is no longer important. The pattern motion does not depend on the segmentation of the scene, but on integration mechanisms and thus also works without feedback of the contrast cells. However, the suppression of extrinsic motion cues along the occluders is crucial to avoid that the strong 2D features generated at the junction of the occluders and the bars are propagated along the bars. This is enabled via the modulatory connections from V2 Form. Also, the barberpole illusion depends on this effect: If occluders are covering the horizontal or vertical line endings, the intrinsic features at these positions become extrinsic features and are as such suppressed by the V2 Form input. As a consequence, the line endings that are not occluded will determine the perceived movement. Here, also the influence of MT integrating cells is important. When the aperture is quadratic, the line endings contribute likewise to a vertical and horizontal percept, the sum of the overall activity results in the normal flow.

In the context of the results of psychophysical and neurophysiological experiments not only the final percept is of interest, but also the temporal change of neural activity leading to different percepts at different times. Here, the solution of the aperture problem comes into play. For both the chopstick and barberpole illusion after the first iteration along the edges of the bars only normal flow is estimated, the true motion can be propagated along the stimulus. The time necessary to complete the propagation of the correct movement increases with bar length, as shown by neurophysiological data of [12] that can be replicated with model areas V1 and MT Motion (as proposed by

[11]). In the case of the chopstick illusion the process of propagation is more difficult: Depending on the configuration, either the activity from line endings or the central pattern motion should propagate (compare Fig. 2 and 3). This means that context information has to be included in the motion processing path in a way that the propagation is steered by the global configuration. In our model, the modulatory input of V2 Form cells and feedback of MT contrast cells to MT Motion represent a “soft switch” that biases the motion areas and allows the correct propagation step by step. The iterative generation of the correct percept is also related to experiments concerning OFRs. The temporal shift of activity from normal flow to pattern flow (as in experiment 2 and 4) is in line with behaviour found for OFR. Kawano et al. [6, 7] measured that the activity in monkey MT/MST is preceding OFR by approx. 10 msec. For experiments showing stimuli first dominated by component and then by pattern motion like the barberpole illusion or moving plaids, experiments both with humans [8, 9] and monkeys [10] revealed that the initial direction of eye movements follow the component motion direction. Over time, the direction changes to the direction of pattern motion. The interaction of form and motion pathway in our model represents a plausible explanation for the generation of different eye movement directions measured in these experiments. The interaction of the model areas provides robust segmentations and disambiguated representations of the scene at a low level of the cortical hierarchy and could generate the necessary representations underlying subsequent processes of sensory-initiated perceptual decision-making.

5. Conclusion

In conclusion, we propose a neural model of early visual processing areas for motion detection, segmentation, and integration that suggests interactions between the motion (V1, MT) and the form (V1, V2) pathway to (i) guide the integration of motion responses in MT, to (ii) avoid salient activity of motion cells at positions with occlusions and (iii) to achieve MT cells tuned to both motion and disparity. We also explain the role of different MT cells, in particular for MT contrast cells that support the segmentation of objects moving in different directions. The different model areas allow a distributed representation of the information which interacts via modulatory connections. We claim that without these connections, the individual areas fail to compute the correct estimate as demonstrated by our lesion experiments. The presented model successfully reproduces experimental data of the chopstick and the barberpole illusion.

Acknowledgements

This research has been supported in part by a grant from the EU (FP6 IST Cognitive Systems, proj. no. 027198).

References

[1] M. Mishkin, L.G. Ungerleider, & K.A. Macko, Object vision and spatial vision: Two central pathways, *Trends in Neuroscience*, 6, 1983, 414–417.

- [2] R.T. Born, D.C. Bradley, Structure and function of visual area MT, *Ann. Rev. Neurosci.*, 28, 2005, 157–189.
- [3] C.R. Ponce, S.G. Lomber, & R.T. Born, Integrating motion and depth via parallel pathways, *Nat. Neurosci.* Jan. 13, 2008 (epub).
- [4] S. Shimojo, G.H. Silverman, & K. Nakayama, Occlusions and the solution to the aperture problem for motion, *Vision Research*, 29, 1989, 619–626.
- [5] S.M. Anstis, Imperceptible intersections: the chopstick illusion (AI and the Eye, John Wiley, New York, USA, 1990).
- [6] K. Kawano, M. Shidara, Y. Watanabe, S. Yamane, Neural activity in cortical area MST of alert monkey during ocular following responses, *J Neurophysiol*, 71, 1994, 2305–2324.
- [7] K. Kawano, Y. Inoue, A. Takemura, T. Kitama, F.A. Miles, A cortically mediated visual stabilization mechanism with ultra-short latency in primates (Parietal lobe contribution to orientation in 3D space, Thier P., Karnath H.-O., Springer, Heidelberg, 1997, 185–199).
- [8] G.S. Masson, L.S. Stone, From Following Edges to Pursuing Objects, *J Neurophysiol*, 88, 2002, 2869–2873.
- [9] G.S. Masson, E. Castet, Parallel Motion Processing for the Initiation of Short-Latency Ocular Following in Humans, *Journal of Neuroscience*, 22, 2002, 5149–5163.
- [10] R. T. Born, C. C. Pack, & R. Zhao, Integration of motion cues for the initiation of smooth pursuit eye movements, *Prog. Brain Res.*, 140, 2002, 225–237
- [11] P. Bayerl, H. Neumann, Disambiguating visual motion through contextual feedback modulation, *Neural Comp.*, 16, 2004, 2041–2066.
- [12] C.C. Pack, R.T. Born, Temporal dynamics of a neural solution to the aperture problem in cortical area MT, *Nature*, 409, 2001, 1040–1042.
- [13] A. Thielscher, H. Neumann, Neural mechanisms of cortico-cortical interaction in texture boundary detection: A modeling approach, *Neuroscience*, 122, 2003, 921–939.
- [14] R. Eckhorn, H.J. Reitboeck, M. Arndt, & P.W. Dicke, Feature-linking via synchronization among distributed assemblies: Simulation of results from cat cortex, *Neural Computation*, 2, 1990, 293–307.
- [15] P. Bayerl, H. Neumann, Disambiguating visual motion by form-motion interaction - a computational model, *IJCV*, 72, 2007, 27–45.
- [16] J. Gautrais, S. Thorpe, Rate coding vs. temporal order coding: A theoretical approach, *BioSystems*, 48, 1998, 57–65.
- [17] E. M. Izhikevich, Simple model of spiking neuron. *IEEE Trans. on Neural Netw.*, 14, 2003, 1569–1572.
- [18] H. Wallach, Über visuell wahrgenommene Bewegungsrichtung, *Psychol. Forsch.*, 20, 1935, 325–380.
- [19] C.C. Pack, A.J. Gartland, & R.T. Born, Integration of contour and terminator signals in visual area MT of alert macaque, *Journal of Neuroscience*, 24, 2004, 3268–3280.
- [20] J. Berzhanskaya, S. Grossberg, & E. Mingolla, Laminar cortical dynamics of visual form and motion interactions during coherent object motion perception, *Spat. Vision*, 20, 2007, 337–395.
- [21] S. Grossberg, The complementary brain: Unifying brain dynamics and modularity, *Trends in Cognitive Sciences*, 4, 2000, 233–246.
- [22] L. Lidén, C.C. Pack, The role of terminators and occlusion in motion integration and segmentation: a neural solution, *Vision Research*, 39, 1999, 3301–3320.
- [23] E. Tlapale, G.S. Masson, & P. Kornprobst, Biological model of motion integration and segmentation based on form cues, *INRIA Rapport de recherche RR-6293*, 2007.