Visual motion disambiguation via contextual feedback control

Olivia Haas*

Supervisors:
Thomas Wachtler
&
Christian Kellner

Lab Rotation
Martinsried (Munich, Germany)
April 25, 2012

Abstract

This lab rotation considered the affect of contextual feedback from the visual area MT to the primary visual cortex V1. Due to the aperture effect the movement of the presented square stimulus is amiguous accept at its edges. A model formulated to disambiguate the movement of each pixel consideres three stages within each model area. The first stage architecture is dominated by modulary feedback from the higher stage, whereas the second stage realizes an integration in space by an isotropic gaussian filter. Finally stage three normalizes estimations to emphasize unambgous signals.

*LMU - Graduate school of Systemic Neurosciences - Munich, Germany

Contents

1 Introcution

In a natrual environment the visual system deals with disambiguating moving stimuli. This process involves area V1 and MT of the dorsal stream, see Figure 1.

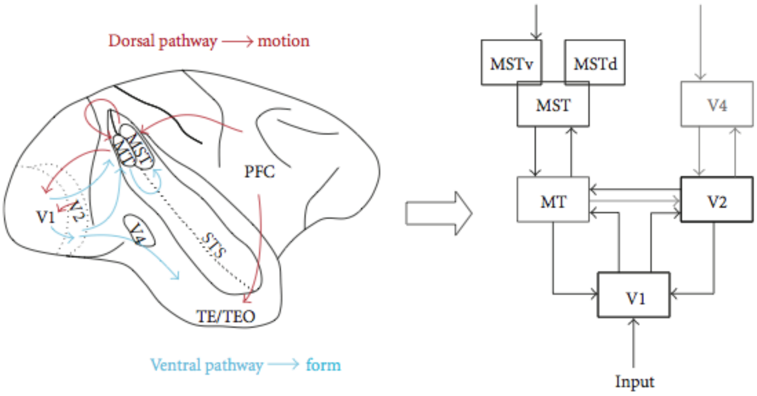

Figure 1: The Dorsal pathway for motion processing. The location of involved brain areas are shown in the left part whereas a structural processing scheme is displayed on the right side.

The input to model area V1 is considered to be the output of a local measure of motion, which is realized by the Reichard detector [?]. It delivers a movement orthogonal to an extended edge, which is known as the aperture problem. This can be resolved for localized image features, such as edges, but remains otherwise. Due to the small receptive field of single neurons a e.g. horizontal edge which is moving in a 45 degree agle seems to move upward instead.
Such an initial detection of motion can be contradicting within one object and remains to be resolved within following areas V1 and MT.
This disambiguation step is further subcategorized in six stages, three within the primary visual cortex and three in MT. Stages of both model areas are equal in their architecture.

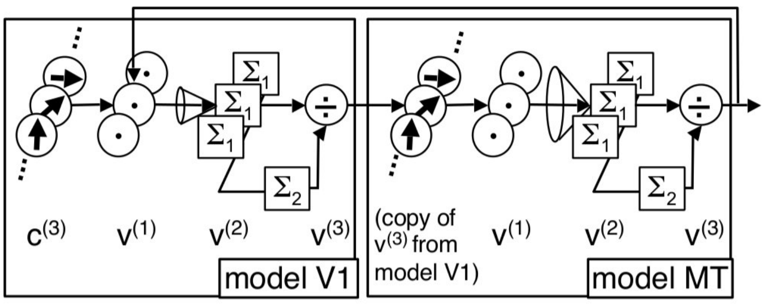

Figure 2: Overview of the model architecture including three stages within area V1 and MT. Feedback stage $v^{(1)}$ is the output of area MT in V1 and zero in area MT, since no higher stages are considered to feed back to area MT. Stage $v^{(2)}$ consists of an integration with isotropic gaussian kernals in space realized by a convolution with the signal from $v^{(1)}$. $v^{(3)}$ is the lateral shunting inhibition stage ???? where unambiguous signals get amplified and amiguous ones get weakened.

In order to solve the aperture problem this architecture involves feedforward and modulatory feedback, which provides a prediction mecanism and is realized via a recurrent loop, see Figure 2. However, in V1 the fead forward signal from the reichard detector receives, as a first modulatory impact, feedback from area MT. The resulting actvity can then be interpreted as a measure of degree of match between the feedback and input at the next time step. Therefore, if the prediction of area MT matches the new input, the resulting activation will be maximal. The prupose of this

step is to provide activations in V1 that match the expactations of the system
Further the resulting signal is sharpened by squaring it and then fead forward to the step of spatial integration. Here the signal from the previous step is convoluted with a two dimensional spatial gaussian. This serves the aim that motion information from neighbouring pixels are smoothened which can be compared to the vector average of the individual motion vectors. Here the model ensures integration of the surround into the motion information of each pixel which is then fead foreward to the final model area.
Here a normalization of movement direction estimations is realized by ?????

2 Theory of Model Stages

2.1 Input Stage

As a stimulus a black square on a white backround was used, which moves in the direction of 45° with a step size of Δt.

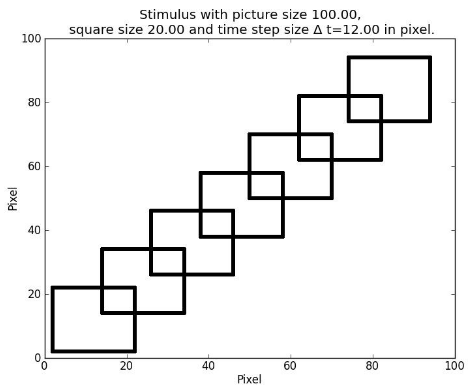

Figure 3: Black square stimulus moving in the direction of 45 °
. Each time step Δt is represented as a single square at a certain position. For illustration reasons in this figure all time steps are presented together.

As described above the input to model area V1 is the output of the Reichard detector. In this lab rotation the reichard detector was not implemented as such but its output was created manually instead.
More explicit this means that for each pixel a population code was build to mimic the Reichard detector output. This was done by choosing 8 neurons which are individually tuned, in a gaussian fashion, to a certain angle in space. The neuronal tuning curves were chosen to equally divide the entire space (360°) and can be seen in figure 4

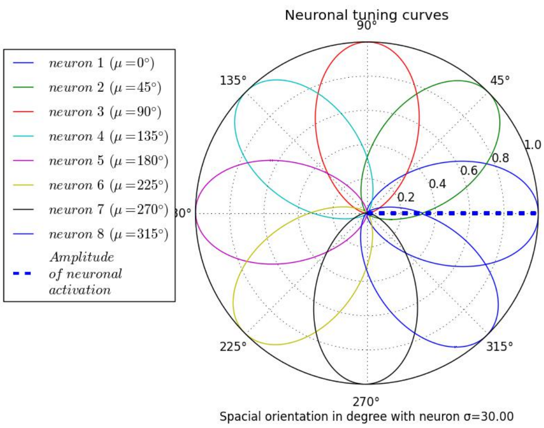

Figure 4: Gaussian tuning curves in space of 8 different neurons with a maximal Amplitude of 1.

For each direction in space one can now write an eight dimensional vector, where the rows correspond to the individual neuron with its activation for the chosen direction in space (see figure 4). Due to the aperture problem we assume vertical vectors for horizontal stimulus edges and horizontal vectors for vertical edges. The three pixels which form the stimulus corners indecate the correct direction of 45°.
In order to plot such vectors the neuronal activation from the population code needs to be translated. This can be done by mutiping the eight dimensional vector containing the activation of the single neurons ($A(n)$) with the corresponding basis vectors:

$$(A(1), A(2), \ldots, A(8)) \cdot \begin{pmatrix} (0 & 1) \\ (1 & 1) \\ \vdots & \\ (-1 & 1) \end{pmatrix} \tag{1}$$