Project on the Early Motion Pathway (Vision)

Cornelia Fermüller

This project looks at the early motion processes. These processes are image motion estimation, 3D motion estimation and motion segmentation. We can find many algorithms in the Computer Vision literature and many models in the biological vision literature. In this project we will study properties of the underlying computations employed by the human vision system to solve these processes. We are going to do this by experimenting with optical illusions. We will create algorithms, and especially neural network implementations that work properly under “normal condition,” but for special patterns (known as illusions) produce the same erroneous prediction as the human perception. In general, the work with illusions should be going both ways. That is, we start with some known illusions to develop our theory. Then, on the basis of our theory we create variants or even new patterns, to test the theory.

An important observation is that the early vision processes are feeding information into each other: The early vision processes, (i.e. image motion estimation, 3D motion estimation, and motion segmentation) are not simple. Looking at these processes from a computational point of view, it is very challenging to obtain very accurate image motion without knowing the boundary of objects and without knowing the 3D motion. On the other hand, to estimate the boundaries of objects, we need good estimates of image motion. Thus, we see that there is a coupling between these processes, and they need to feed information to each other and be solved iteratively. On the other hand, the large amount of feedback we find in the biological system confirms these ideas.

There are two main aspects we will explore.

Replicating human perception

One aspect is to study properties neural networks. We can look at classical networks and spiking networks. We would like to understand the properties that these networks need to have (architectures, cost functions, ..) to produce results as in the perception of illusions. Specifically, we will start with two kinds of patterns: The Snake patterns and the Ouchi pattern.

In previous works, we hypothesized, that the main reason for these illusions lie in the way image motion (or optical flow) is computed. In the case of the Snake pattern, it is “causal optical flow, “as explained in this paper. The vision system uses temporal filters that compute flow from the present and the past with more weight giving to the present. In the case of the Ouchi patter, the reason is statistical bias, as explained in these papers (1, 2), and illustrated at this website. To estimate image motion, one needs to solve an estimation problem. However, with noise, the estimation is statistically biased, and thus on a pattern, with only two texture directions, but many more in one direction than the other, there will be erroneous flow estimation. We can experiment with networks for flow architectures, and compare them. In both illusions, the snake and the Ouchi, image motion is created by the small eye movements during fixation. These eye movements are usually discarded: Our vision system seems to compute the rotation of the eye from the images. However, in the case of the illusions, there is residual motion, which is perceived as independent movement. Besides the flow, there also is a clear segmentation of parts of the pattern, and even a depth perception in the case of the Ouchi pattern. We also will look into the interplay of segmentation and flow.

Code to start with has been set up at https://github.com/prgumd/motion_illusions

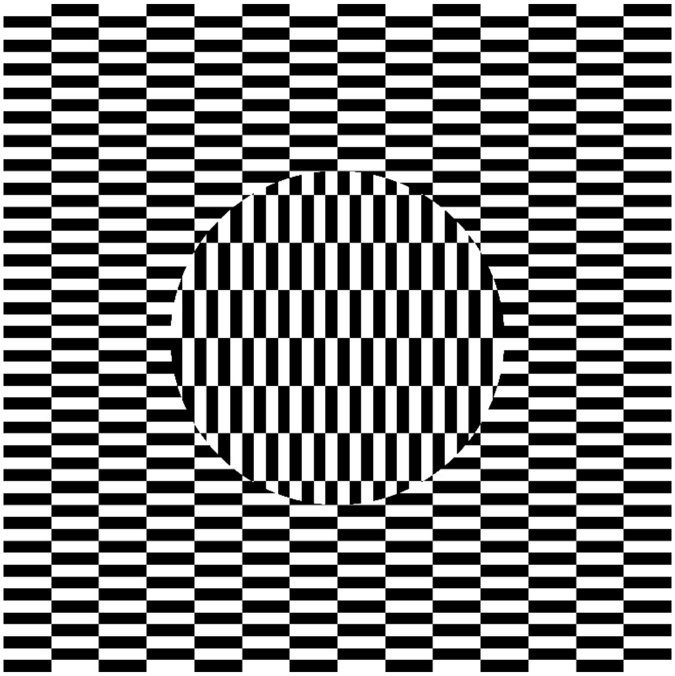
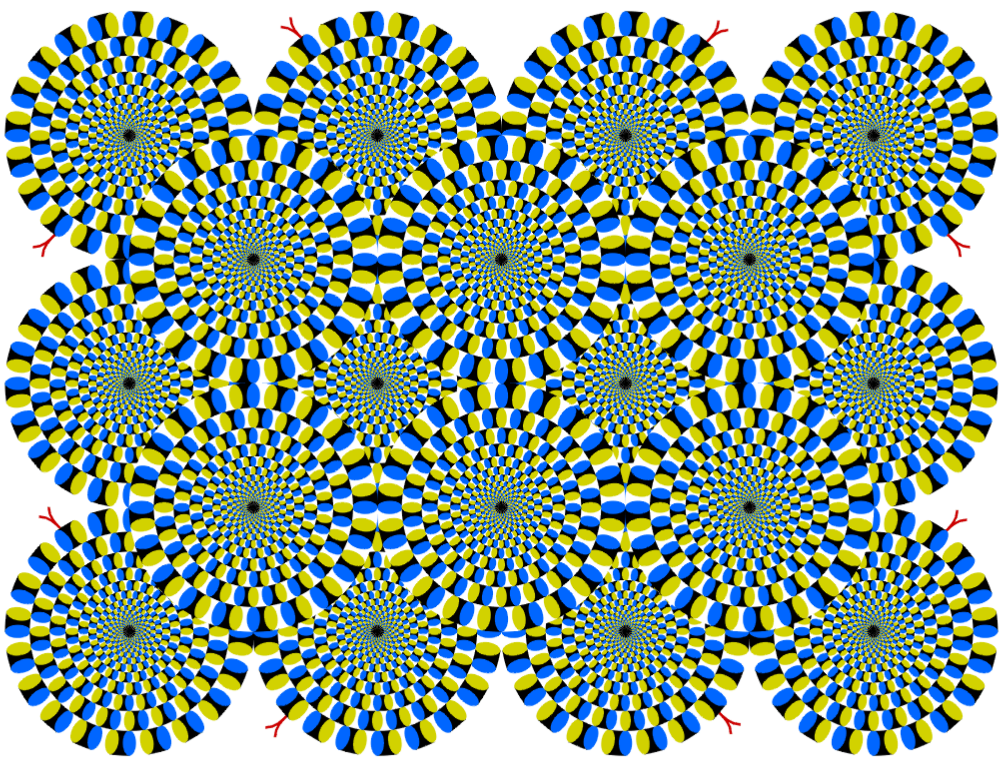


Figure 1: Ouchi Illusion. Small retinal motions, or slight movements of the paper, cause a segmentation of the inset pattern, and motion of the inset relative to the surround.



.

Figure 2: Rotating Snake. In peripheral vision most observer experience rotary movement in both patterns. The direction in the circular arrangements alternates, with counter-clockwise direction in the upper left.

The Role of event signals

The second aspect is about looking into the role of the transient signal. We strive to look at both human vision and computational aspects for technology. Our hypothesis is that the event signal brings advantages for motion segmentation of moving objects. There may (should) be others. To start with, we will look into the stepping feet illusion. We will create variants (as in Fig. 3) to explore the role of contrast and events in the perception of this signal. For information on retinal ganglion cells and their role in computing sustained and transient response, see here. Our hypothesis is that this illusion does not happen simply because of a process that matches the moving rectangles or because of flow, but because events are recorded and processed, depending on the contrast. We can build models for the flow estimation and segmentation and show their implementation in classic and neural networks.

We hope to discuss other aspects and situations, where we can establish that events give us a clear advantage over classic frames in computational processes. We can look at tracking, challenging segmentation scenarios, and the interplay of different cues in grouping. We also should look at the advantages that we gain by using in geometric computations early on only motion information, as opposed to all the image information, which leads to overfitting.

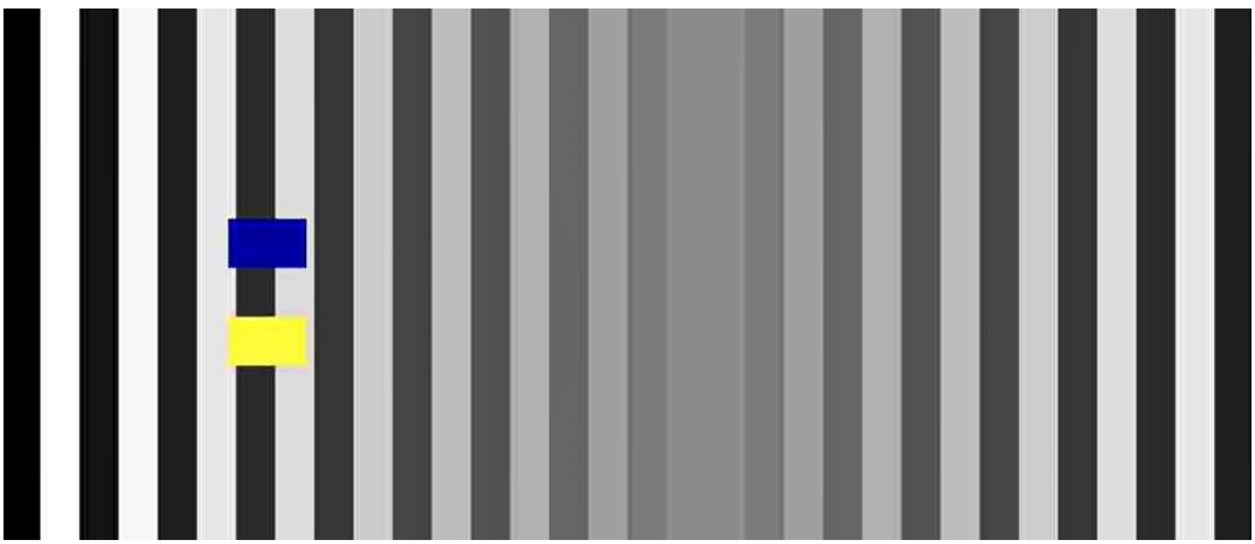


Figure 3: A variant of the Stepping Feet illusion to study the impact of contrast and events.