Summary

Individual cells in the retina respond to only a small portion of the visual scene and thereby send a fragmented representation of the outside world to the rest of the visual system. The visual system transforms this representation into a coherent percept of the visual scene in which objects are perceived as being in front of a background. This process is termed *perceptual grouping*. This PhD thesis presents experiments that aim to enhance our understanding of the neural basis of perceptual grouping in rhesus macaques and humans.

Each neuron the the primary visual cortex responds to a small area of the visual scene. This area is the receptive field of the neuron. A neuron with a receptive field that overlaps with a figure fires action potentials at a higher rate than neurons with a receptive field on a background. The difference in firing rate is known as figure-ground modulation (FGM). FGM in early visual cortex is thought to arise due to feedback from higher visual areas. FGM has typically been measured using relatively small, texture-defined squares on a uniform background in which information only needs to be integrated over a small spatial scale. The possibility therefore remains that FGM arises through local computations in lower visual areas instead of feedback from higher visual areas. The first study described in this PhD thesis investigates figure-ground perception in humans and macaques and the neural responses in V1 and V4 in the macaque to stimuli in which figure-ground relationships are defined by Gestalt laws of 'symmetry', 'convexity', and 'enclosure'. Because the local features in and around the receptive fields of the neurons are matched for figure and background regions, visual information needs to be integrated over a large spatial scale to establish a figure-ground modulation. We first investigated whether the segregation of a Gestalt stimulus occurred by measuring the perceived contrast of an embedded Gabor patch in the figure or background regions. The behavioral results showed that both humans and monkeys indeed perceived Gabors on figures as higher in contrast than Gabors on the background, indicating a successful figure-ground segmentation. The neural responses recorded from the monkey shows figure-ground modulation for Gestalt figures in V1 as well as in V4. The strength of this modulation correlated with the strength of perceptual segregation as judged by the perceived contrast differences. We conclude from this that neurons in early visual areas are sensitive to information that is far outside of their classic receptive field. This result implies that figure-ground modulation in early visual cortex is established through feedback from higher visual areas.

The second study investigates the role of learning and attention on FGM. The visual stimuli in this experiment consist of 'proto-objects', which are regions of the visual scene that have many of the statistical properties of objects. To correctly segment figures from their ground, responses to proto-objects that belong to the figure need to be enhanced and responses to proto-objects

that belong to the ground need to be suppressed. We recorded neural activity from V1 and V4 of macaque monkeys while they discriminated between N- and U-shaped forms. Activity in V1 showed FGM within 90 ms of stimulus onset. A suppression of the proto-object belonging to the ground was present in naive animals. This suppression became stronger over the period in which the animal learned the task. The enhanced activity for the figure correlated with performance on the shape discrimination task. In a task in which the animal attended to a second shape on the opposite side of the screen, we found that figure-ground modulation is weakened for the unattended figure in the receptive fields of the recorded neurons. Cells in V4 showed selectivity for particular shapes or particular borders of the figures in the discrimination task. For example, some cells preferred the N-shape over the U-shape, and other cells responded more to the outer edge than to the inner edge of the shape. The figure-ground effects observed in V1 may arise through feedback connections from cells representing specific shapes and boundaries in areas such as V4, where FGM occurred within 70 ms.

The third study described in this thesis reveals the important role of attention in perceptual grouping. After the initial figure-ground structure is established in the visual scene, an object can be selected by attention to complete the perceptual grouping process. We investigated the time course of perceptual grouping of two-dimensional objects and showed that it is associated with the gradual spread of object-based attention across the object's surface. We compare several neurocomputational models that aim to explain the time course of attentional selection. The predictions of the models are compared to the behavioral data from human participants in three separate behavioral experiments. The data shows that attention spreads fastest over large and homogeneous areas and is slowed down at locations that require small-scale processing. The growth-cone model accounted best for the observed data. This model takes into account the receptive fields size of the cells in various visual areas.

The fourth study suggests an explanation for a curious finding we stumbled upon in the third study: the growth-cone model was in excellent agreement with the behavioral data for objects that were simple or unfamiliar to the participant, but its predictive value decreased for familiar objects in natural scenes. We hypothesized that the brain can take advantage from the extensive visual experience we accumulate throughout our life with a wide variety of objects. We tested the idea that perceptual grouping for familiar stimuli such as animals and vehicles may benefit from selectivity in cells in higher visual areas to parts of an object that can provide feedback to cells in lower visual areas involved in perceptual grouping. Participants could very fast and efficiently classify images as animals or vehicles in a classification task. In an image-parsing task, participants reported whether two cues fell on the same or different objects, and we measured reaction times. Despite the fast object classification, perceptual grouping required more time if the distance between cues was larger, and we observed an additional delay when the cues fell on different parts of an object. Parsing was slower for objects in an unfamiliar orientation (upside down) than for upright objects. The data shows that for familiar objects, the time course of perceptual grouping is affected by the knowledge we have about the common structure of the object. The results imply that perception starts with rapid object classification that is followed by a serial perceptual grouping phase.