Abstract Dr. Jan Tünnermann

Spatiotemporal Preattentional Structures: A New Approach to Articial Visual Attention

Abstract:
Vision is an important source of information for humans and other animals. Their ability to effortlessly select and process relevant and filter out irrelevant information from the abundant and high-dimensional raw input has intrigued developers of technical cognitive systems for a long time. For more than thirty years, various methods have been put forward to transfer these selective attention mechanisms to technical systems in computer vision and robotics. In this field, two major trends can be identified. In one line of research approaches parochially focus on mimicking human attention. This may often not be justified, because the technical target systems share neither the exact tasks, environments, and embodiments, nor the evolutionary and developmental histories of their biological counterparts. Approaches in the second trend have diverged in opposite direction: They neglect much of the original concepts of selective attention as known from biology and have reduced the task to extracting “saliency” (conspicuous elements) from images with arbitrary methods. Both extremes hamper the use of artificial attention in computer vision for autonomous agents as, for example, mobile robots: For the close-to-biology approaches, interfacing with other cognitive modules can be a challenge. The technical representations of knowledge, current tasks, and intentions, which are known to influence attention, are typically heterogeneous and different from those in biological systems. Similarly, it is often unclear how to connect such attention models to robotic actions beyond tuning camera heads. For the models that compute saliency by virtually arbitrary methods, it is even less clear how they can be combined with top-down influences, how dynamic scenes can be processed, or how their results can be turned into behavior.

The approach to artificial attention introduced in this presentation is a remedy for these issues. It uses the novel concept of “spatiotemporal preattentional structures”, that is, basic abstract elements that are extracted early in the attention process and that describe the input in space and time. Their extraction itself does not closely follow biological visual processing, but once theyareavailable, themechanismknownfromhumanvisualattentioncanbe applied on a convenient abstract level that integrates more seamlessly with other technical cognitive modules. On this basis, important advances in the artificialattentiondomaincouldbereached. Amongthemarethecalculation of spatiotemporal saliency, attentional orienting not only in space but also in time, a good performance in salient-object detection without sacrificing basic attention concepts, integrating object affordances (action possibilities) in the control of attention, and a “selection-for-action” framework to move toward a more behavior-centered concept of artificial selective attention.