Startseite
Programm
Anmelden
Beitrag einreichen
Poster-Info
Tübingen-Info
Frühere Konferenzen
Veranstalter
Kontakt

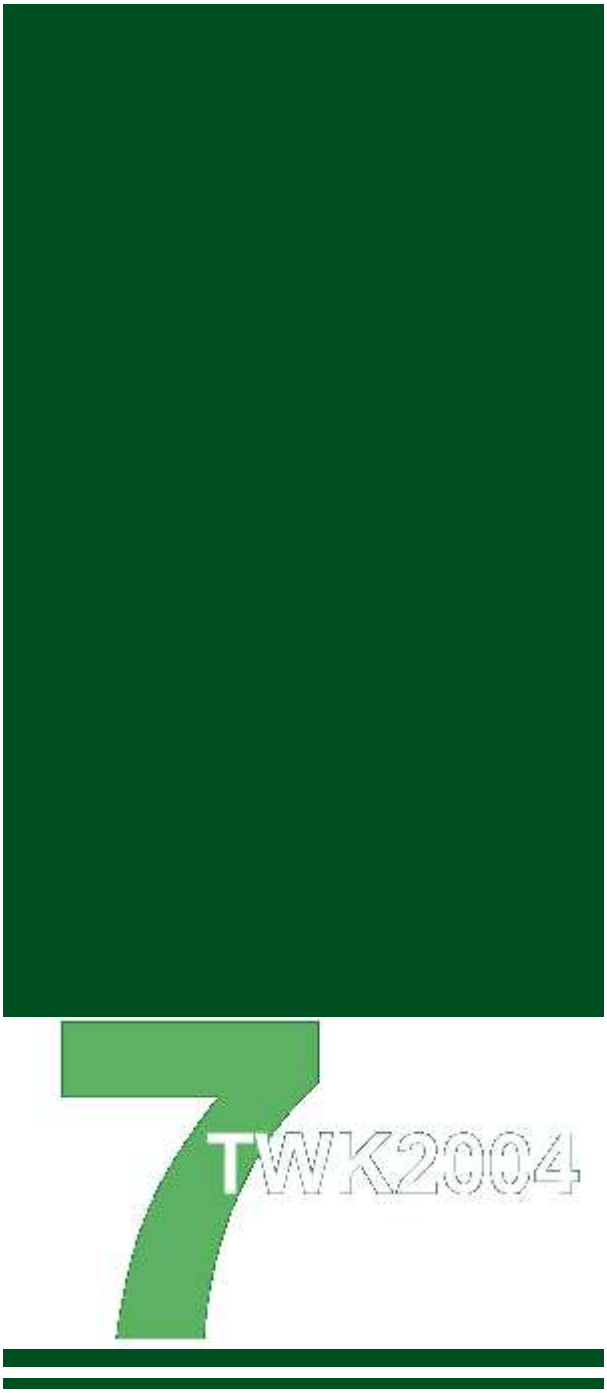
Focussing Object Recognition on Regions of Interest

Simone Frintrop, Andreas Nüchter & Hartmut Surmann (Fraunhofer Institut AIS, St. Augustin)

While humans are able to detect and recognize objects in their environment effordlessly, these tasks belong to the hardest problems in computer vision. One of the difficulties lies in the huge amount of data that has to be processed. In this work we restrict object classification to regions of interest detected by biologically motivated attention mechanisms. This yields a significant increase in speed of classification.

The sensor data used for our experiments are acquired by a 3D laser scanner mounted on top of an autonomous mobile robot. The scanner yields illumination independent range and reflection data, both visualized as gray-scale images. The advantage over camera images is, firstly, the illumination independence enabling the acquisition of the same data in sunshine as in complete darkness and secondly, the precise and dense depth information that is only achievable from camera data with difficulties.

The laser images are fed into a biologically motivated attention system. In human vision, attention helps identify relevant parts of a scene and focus processing on corresponding sensory input. Inspired by the psychological work of Treisman [1], computational attention models determine saliencies according to different features like intensity, color, and orientations in parallel and fuse them into a single map. The most salient region in this map determines the focus of attention. In contrast to other models, our approach fuses saliencies of both laser modalities, depth and reflection, to achieve a single focus of attention. The fusion of sensor modalities is performed in



analogy to humans who use information from all senses and enables considering more object properties.

The regions of interest detected by the attention system serve as starting points for a fast cascade of classifiers [2]. In contrast to the conventional approach, we search for a previously learned object only near the focus of attention instead of considering the whole image. This enables a concentration on relevant image regions and a significantly faster classification, requiring only 30% of the time of an exhaustive search (60ms vs. 200ms, Pentium-IV-2400 MHz). Considering also the time for the attention system (230ms), the system performance already increases for two learned object classes and the time saving increases proportionally with the number of object classes.

In future work, the recognized objects will be registered in semantic 3D maps that are automatically created by the autonomous robot. The maps will support self-localization of the robot and serve as an interface between robot and humans.

[1] Treisman, A. and Gelade, G .(1980): Cog. Psych. 12, 97-136
[2] Viola, P. and Jones, M. (2001): Proc. 2nd Int'l Workshop on Statistical and Computational Theories of Vision - Modeling, Learning, Computing and Sampling