Research

Research at IIS is situated at the intersection of computer vision, machine learning and robotics, and focuses on adaptive perception-action systems as well as on image and video analysis. Some of our areas of activity include

- object models for robotic interaction
- perception for grasping and manipulation
- systems that improve their performance with experience
- video analysis for applications such as sports or human-computer interaction
- links with the psychology and biology of perception

Most of our work is based on probabilistic models and inference.

The following list highlights some examples of our current and past work. Follow the links for more information.

Computational neuroscience - Object recognition is a very hard problem, the best proof of this is to consider that the first works started in the 1960s (Robert's and Guzman's theses) and up to now there are thousands of papers published each year with solutions to achieve that goal. Due to the complexity of this problem, many scientists and engineers have resorted towards neurophysiology for solutions on how the human visual system solves such a difficult task with astonishing efficiency and accuracy. The earliest models inspired by the primate visual system appeared little after the influential works of Hubel and Wiesel (1962,1965,1968) that revealed key mechanisms of the functional architecture of the visual cortex. Computational neuroscience deals with the simulation of the different functions of biological neurons by approximating that functionality by means of mathematical approximations. The recent approaches to object recognition are mainly driven by the `edge doctrine` of visual processing pioneered by Hubel and Wiesel's work. Edges, as detected by responses of simple and complex cells, provide important information about the presence of shapes in visual scenes. We consider that their detection, however, is only a first step at generating an interpretation of images that is invariant against certain variances in scene properties and layout as well as their geometric transformations. Thus, our line of work is on intermediate processing. Intermediate processing areas operate upon initial simple and complex cell outputs lead to the formation of neural representations of scene content that allow robust shape inference and object recognition. We focus on the steps that computer models in our opinion must pursue in order to develop robust recognition mechanisms that mimic biological processing capabilities beyond the level of cells with classical simple and complex receptive field response properties. Check our publications Rodriguez-Sanchez and Tsotsos, 2012 and Rodriguez-Sanchez et al, 2014.

Multimodal, hierarchical models for object manipulation - We have developed a generic framework for object grasping using our robotic platform. The main software modules perform object detection and pose estimation, grasp planning, path planning and robot arm and hand trajectory execution. Over the course of the PaCMan project, we will replace many of these modules by advanced modules developed by the project. For example, the object detection component, currently based on state-of-the-art methods, will benefit from hierarchical compositional models based on both 2D and 3D information. This representation will be useful both for reasoning in a cluttered

environment where only parts of an object will be visible, and for improving object manipulation.

Probabilistic models of appearance for object recognition and pose estimation in 2D images - We developed methods to represent the appearance of objects, and associated inference methods to identify them in images of cluttered scenes. The goal here is to leverage, to a maximum, the information conveyed by 2D images alone, without resorting to stereo or other 3D sensing techniques. We are also interested in recovering the precise pose (3D orientation) of objects, so as to ultimately use such information in the context of robotic interaction and grasping.
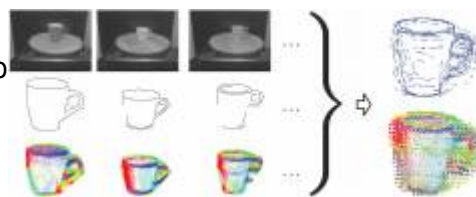
Grasp Densities: **Dense, Probabilistic Grasp Affordance Representations** - Motivated by autonomous robots that need to acquire object manipulation skills on the fly, we are developing methods for representing affordances by nonparametrically by samples from an underlying affordance distribution. We started by representing graspability by distributions of object-relative gripper pose distributions for successful grasps (Detry et al. 2011).

Probabilistic Structural Object Models for Recognition and 3D Pose Estimation - Motivated by automnomous robots that need to acquire object models and manipulation skills on the fly, we developed learnable object models that represent objects as Markov networks, where nodes represent feature types, and arcs represent spatial relations (Detry et al. 2009). These models can handle deformations, occlusion and clutter. Object detection, recognition and pose estimation are solved using classical methods of probabilistic inference.

Reinforcement Learning of Visual Classes - Using learning approaches on visual input is a challenge because of the high dimensionality of the raw pixel data. Introducing concepts from appearance-based computer vision to reinforcement learning, our RLVC algorithm (Jodogne & Piater 2007) initially treats the visual input space as a single, perceptually aliased state, which is then iteratively split on local visual features, forming a decision tree. In this way, perceptual learning and policy learning are interleaved, and the system learns to focus its attention on relevant visual features.
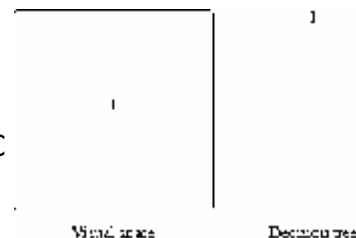
Reinforcement Learning of Perception-Action Categories - Our RLJC algorithm (Jodogne & Piater 2006), extends RLVC to the combined perception-action space. This constitutes a promising new approach to the age-old problem of applying reinforcement learning to high-dimensional and/or continuous action spaces.
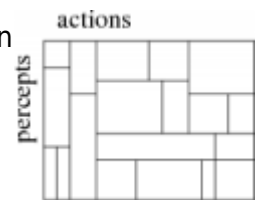
Image Classification with Extremely Randomized Trees - We developed a generic method that achieves highly competitive results on several, very different data sets (Marée et.al. 2005). It is based on three straightforward insights: *randomness* to keep classifier bias down, *local patches* to increase robustness to partial occlusions and global phenomena such as viewpoint changes, and *normalization* to achieve invariance to various transformations. The key contribution was probably the demonstration of how far randomization can take us: Local patches are extracted at random, rotational invariance is obtained by randomly rotating the training patches, and classification is done using Extremely Randomized Trees.
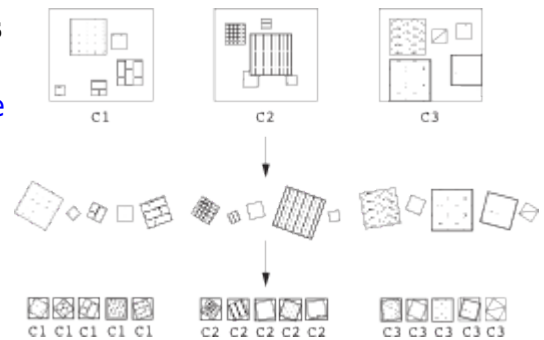
Real-Time Object Tracking in Complex Scenes - While most work at the time was based on background subtraction, we developed new methods for complex scenes by tracking local features for robustness to occlusions and to background changes, taking spatial coherence into account for robustness to overlapping, similar-looking targets. In the context of soccer, we also developed methods for robust, model-based and model-free, absolute and incremental terrain tracking.

From:
https://iis.uibk.ac.at/ - **IIS**

Permanent link:
https://iis.uibk.ac.at/research?rev=1406648678

Last update: **2018/09/03 14:57**