TECHNISCHE UNIVERSITÄT MÜNCHEN
LEHRSTUHL FÜR STEUERUNGS- UND REGELUNGSTECHNIK
UNIV.-PROF. DONGHEUI LEE

June 17, 2016

M A S T E R ' S T H E S I S
for
Tobias Watzel
Student ID , Degree

Learning to Grasp in Highly Cluttered Environments Leveraging Multimodal Perception and Deep Neural Networks

Problem description:

Grasping objects in unstructured and cluttered environments is an open problem in robotics. In the last decades, researchers in the robotics community have proposed different approaches. On one side, classical planning methods require the complete knowledge of the object to grasp and of the environment. On the other hand, data-driven methods have been developed, which exploit machine learning techniques to compensate lack of knowledge and to improve incrementally the performance. Recently, in the context of data-driven techniques, approaches based on Convolutional Neural Networks (CNNs) have been proposed [1]. It has been showed within the machine learning community that deep learning methods shine in presence of a very complex problems and a huge amount of data. These features are potentially suitable in problems of grasping in presence of cluttered scenes where several objects overlap and the models of the objects are unknown. There are only few methods based on CNNs proposed in the literature and they exploit only images. The objective of the thesis is to design a deep neural network that leverage multimodal perception data such as tactile, force, and visual information to learn the connection between multimodal perception and action. In more detail the network will approximate the function that connect visuo-tactile perception with the joint velocities in order to perform successful grasps. In robotic application, collect such a high number of example is not trivial. The effectiveness of simulation methods to bootstrap the training process will be investigated.

Tasks:

- Research the literature on deep networks and in particular CNNs
- Set-up the simulation environment to collect the training data
- Starting from the architecture in [1], develop an architecture that exploits multimodal perception
- Show experimentally the performance of the developed architecture

Supervisor:	Pietro Falco
Start:	XX.XX.2015
Intermediate Report:	XX.XX.2015
Delivery:	XX.XX.2015

(D. Lee)
Univ.-Professor

Bibliography:

[1] Sergey Levine, Peter Pastor, Alex Krizhevsky, and Deirdre Quillen. Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection. *arXiv preprint arXiv:1603.02199*, 2016.