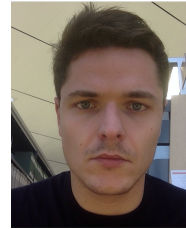


3D Object Recognition with Deep Networks

3D Photography Project Proposal
Supervised by: Moritz Oswald, Pablo Speciale
March 7, 2015

GROUP MEMBERS

Tobias Grundmann Adrian Schneuwly Johannes Oswald



I. DESCRIPTION OF THE PROJECT

The main goal of our project is to enhance object recognition by making use of the recent availability of 2.5D data through Microsoft Kinect, Google Project Tango, et al. We will follow recent approaches [3], [2] which use convolutional deep neural/belief networks to learn the distribution of complex 3D shapes across different object categories of CAD models. Given 2.5D data of a single view point, we will try to recognise the objects category after an intense training of our deep network on a large scale 3D CAD model dataset. Fig. 1

II. WORK PACKAGES AND TIMELINE

A. Prerequisites in March

The first goal of the team is to fully understand the approaches described in [3], [2]. Therefore the team will work through the Udacity deep learning course [1] and extract knowledge and hands-on experience with deep networks for object recognition on 2D data. At the end of the month, the team will be able to understand Deep Networks, its machine learning theory and TensorFlow (in Python) in order to develop a game plan for the project. The division of work between the 3 project members will then be decided.

B. Data Preparation & Modeling in April

The first task is to build a data sets for testing and training our deep network. Therefore we need to extract 2.5D single view point data from our large 3D Cad model dataset. This will be done for the a 40-Class Dataset provided by [3]. Afterwards we can begin reimplementing the papers approaches in Python by starting with the deep learning udacity courses framework. [1] Therefore we will remodel the given 2D object recognition network into a rotation-invariant 3D recognition network which makes heavily use of convolution and pooling.

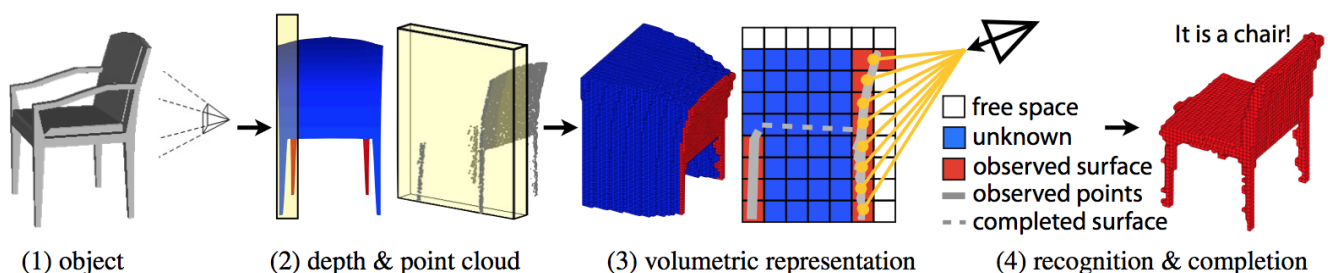


Fig. 1.

C. Training & Testing in June

After successful implementation and testing, we will train the deep network on the hole 40 classes (optional: [3] offers an even larger data set) on own or universities GPU's. Due to lack of experience, we expect a long training phase since the authors algorithms took days to complete.

III. OUTCOMES AND DEMONSTRATION

Goal of the project is successfully reimplement the papers approaches for 3D Object Recognition and achieve very similar positive object recognition results. In our live demo, we will snapshot random CAD Models and try to recognise the object through our algorithm.

REFERENCES

- [1] <https://www.udacity.com/course/deep-learning-ud730>.
- [2] D. Maturana and S. Scherer. Voxnet: A 3d convolutional neural network for real-time object recognition. *International Conference on Intelligent Robots and Systems (IROS2015)*, 2015.
- [3] A. Khosla F. Yu L. Zhang X. Tang Z. Wu, S. Song and J. Xiao. 3d shapenets: A deep representation for volumetric shapes. *Proceedings of 28th IEEE Conference on Computer Vision and Pattern Recognition (CVPR2015)*, 2015.