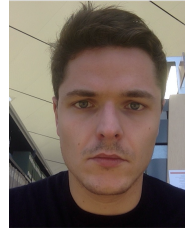


3D Object Recognition with Deep Networks

3D Photography Project Proposal
Supervised by: Martin 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

Deep Networks heavily depend on big datasets. Therefore we are going to use the ModelNet dataset from [3], a CAD dataset of 660 objects. The datasets will transformed into 32x32x32 voxel data and split into a training, valuation and test set. If possible we will use data from the project tango tablet to cross test. We begin reimplementing the VoxNet Deep Convolutional Neural Network 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. The output of the deep network is the the class of object the network thinks it recognized from the input.

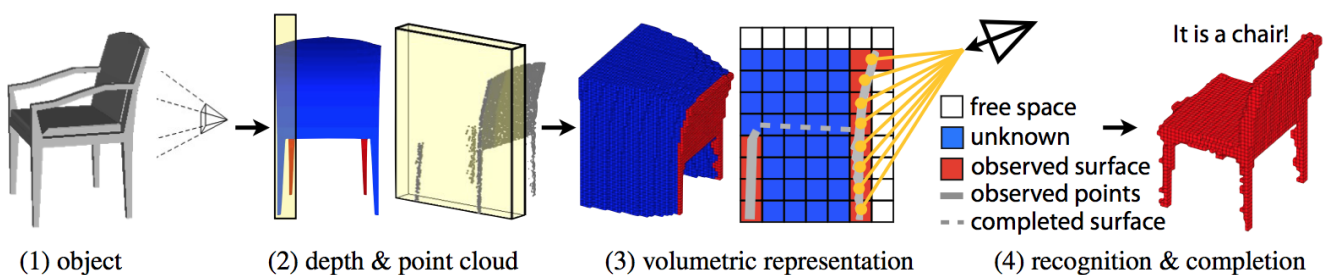


Fig. 1.

C. Training & Testing in June

After successful implementation and testing, we will train the deep network on a subset of 40 classes called the ModelNet40 (optional: [3] offers an even larger data set) on own or universities GPU's. Due to the computationally expensive training of the Deep Network on the large CAD dataset, we will try to start as soon as possible with this phase of the project.

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] Z. Wu, S. Song, A. Khosla, L. Zhang F. Yu, X. Tang, and J. Xiao. 3d shapenets: A deep representation for volumetric shapes. *Proceedings of 28th IEEE Conference on Computer Vision and Pattern Recognition (CVPR2015)*, 2015.