

# Online 3D Object Recognition

Luís A. Alexandre

July 2018

## 1 Abstract

There is a recent interest in using 3D data given the availability of cheap RGB-D sensors [8]. Object recognition is among the many different tasks that the information from these sensors can be used for [1].

The usual approaches to learning consider a setup where the learning is made offline before the classifier is applied to the tasks.

In this thesis we want to develop a online learning approach to solve 3D object recognition problems, expanding on the work in [2], introducing the capability of functioning online with continuous learning.

## 2 Goals

The main goal consists in developing a spatio-temporal filter [3–7, 9] capable of extracting object representations that can be used to recognize objects and their poses from 3D images (point clouds), and making it work online.

Another goal consists in making the code available to other researchers.

The code is developed in linux using C/C++. The PCL [8] will be use for 3D point cloud capture and processing.

A good knowledge of programming, probabilities and English language is necessary.

## 3 Tasks

T1: State-of-the-art in 3D object recognition and spatio-temporal learning methods.

T2: Evaluate and try to adapt existing approaches for different problems (activity recognition, face recognition, etc.) to our problem.

T3: Implement new spatio-temporal filters for object recognition taking advantage of TL approaches.

T4: Make extensive evaluation on public data sets.

T5: Write the thesis and a scientific paper.

## 4 Schedule

<b>Task</b>	<b>Start date</b>	<b>Duration</b>
T1	2018-10-01	2 months
T2	2018-12-01	2 months
T3	2019-02-01	1 month
T4	2019-03-01	1 month
T5	2019-04-01	3 months

## References

- [1] Luís A. Alexandre. 3D descriptors for object and category recognition: a comparative evaluation. In *Workshop on Color-Depth Camera Fusion in Robotics at the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Vilamoura, Portugal, October 2012.
- [2] Luís A. Alexandre. 3D object recognition using convolutional neural networks with transfer learning between input channels. In *13th International Conference on Intelligent Autonomous Systems*, volume 301 of *Advances in Intelligent Systems and Computing Series*, Padova, Italy, July 2014. Springer.
- [3] J. Hirel, P. Gaussier, M. Quoy, J. P. Banquet, E. Save, and B. Poucet. The hippocampo-cortical loop: Spatio-temporal learning and goal-oriented planning in navigation. *Neural Networks*, 43:8–21, July 2013.
- [4] Hema S. Koppula and Ashutosh Saxena. Learning spatio-temporal structure from rgb-d videos for human activity detection and anticipation. In *International Conference on Machine Learning*, 2013.

- [5] Wei Liu, Zhifeng Li, and Xiaoou Tang. Spatio-temporal embedding for statistical face recognition from video. In *Proceedings of the 9th European conference on Computer Vision - Volume Part II, ECCV'06*, pages 374–388, Berlin, Heidelberg, 2006. Springer-Verlag.
- [6] Marianna Madry, Liefeng Bo, Danica Kragic, and Dieter Fox. St-hmp: Un-supervised spatio-temporal feature learning for tactile data. In *IEEE International Conference on Robotics and Automation (ICRA)*, May 2014.
- [7] Akitsugu Noguchi and Keiji Yanai. A surf-based spatio-temporal feature for feature-fusion-based action recognition. In *Proceedings of the 11th European conference on Trends and Topics in Computer Vision - Volume Part I*, pages 153–167, Berlin, Heidelberg, 2012. Springer-Verlag.
- [8] R.B. Rusu and S. Cousins. 3D is here: Point Cloud Library (PCL). In *IEEE International Conference on Robotics and Automation (ICRA)*, Shanghai, China, May 2011.
- [9] Yuke Zhu, Tian Lan, Yijian Yang, Steven N. Robinovitch, and Greg Mori. Latent spatio-temporal models for action localization and recognition in nursing home surveillance video. In *IAPR Conference on Machine Vision Applications (MVA)*, 2013.