

UNIVERSITÀ DI PISA



Department of Computer Science
Degree course in Computer Science

Machine learning techniques for 3D point clouds

Relatori:

Prof. Davide Bacciu

Prof. Daniela Giorgi

Prof. Paolo Cignioni

Prof. Francesco Banterle

Presented by:

Mattia Angelini

**Winter Session
Academic Year 2018/2019**

Contents

1	Introduction	2
1.1	Other 3D techniques	2
1.2	Multi-views	2
1.3	Voxel	2
1.4	Surface-based	3
1.5	Learning on graphs	3
2	Background and Related Works	3
2.1	3D Point Clouds	4
2.1.1	Pointnet	4
3	Techniques for 3D Point Clouds	4
3.1	Sampling	4
4	Results	4
5	Observations	4
6	Conclusions	4

1 Introduction

3D data can be represented in many different ways, where each one inspired new approaches to Deep Learning techniques with their pro and cons.

However in the last years, the importance of 3D point clouds have risen consistently, due to the use in the sensor of smartphones and self-driving cars, but despite that, the focus remains only on how to process those data improving on the models themselves, without put efforts in how to encode those data.

We then have chosen to direct our efforts, on how to improve on the encoding of the information on 3D Point Clouds and, see how this reflects on the results of the currently developed models.

1.1 Other 3D techniques

Before proceeding, we would like to introduce other machine learning techniques on 3D data, starting with the

1.2 Multi-views

Multi-views systems use a group of view of a 3D object to capture its spatial information enabling us to use the success achieved in the Deep Learning for images directly to the views. The main focus here is to find a way to obtain those views, and how feed those view to the machine learning algorithm, the MVCC[6] newtworks was one of the first using this technique, suffering from a big memory footprint.

1.3 Voxel

Voxels on the other hand, encode the information of the 3D data in a regular 3-dimensional grid, where each cube represent a unit of volume in which the 3D information is stored, giving us a volumetric representation of the object. Having a 3D grid means that we could extend the techniques like the convolution to this data rapresentation[5], at the cost of computational complexity passing by $O(n^2)$ to $O(n^3)$ due to the new dimension. New techniques improve on

this using a subdivision of the grid in OctTrees with patches to aggregate groups of Voxels[8], using this method is possible to work on bigger and more detailed grids. Voxels are also used to process point cloud data, here the points are voxelized, enabling the use of voxel techniques to process point cloud data sacrificing some information and accuracy about the object surface, expressed from 3D point clouds.

1.4 Surface-based

Surface-based representation, aims to understand the information of objects using their superficial information, however, working with this type of data is not straightforward due to their non-uniform representation. To overcome this issue many solutions are proposed, from work on the 3d using special convolution[3] to use Geodesic patch to encapsulate the shapes of the surface in images to pass through a CNN network[4]. However, surface-based approaches usually require well structured meshes used to represent the surface, making it difficult to work in real world scenarios with noise.

1.5 Learning on graphs

Machine learning applied to graph is a different field from the 3d based learning sector, despite that we can find some similarity from the data structure of graphs and meshes as well as cloud points. It is interesting how the core framework used in machine learning on graphs, the Message Passing Neural Network[1], can be adapted to perform the same procedure in some scenarios of the 3D based machine learning like the before mentioned MeshCNN[] or the PointNet++[]. Despite the new state of the art techniques like the Deep Graph Infomax[7], and the GraphSage[2] that improve on the core technology have still to be implemented for 3D data.

2 Background and Related Works

In the last years 3D point clouds

2.1 3D Point Clouds

2.1.1 Pointnet

3 Techniques for 3D Point Clouds

3.1 Sampling

4 Results

5 Observations

6 Conclusions

References

- [1] Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, and George E. Dahl. Neural message passing for quantum chemistry, 2017.
- [2] William L. Hamilton, Rex Ying, and Jure Leskovec. Inductive representation learning on large graphs, 2017.
- [3] Rana Hanocka, Amir Hertz, Noa Fish, Raja Giryes, Shachar Fleishman, and Daniel Cohen-Or. Meshcnn. *ACM Transactions on Graphics*, 38(4):1–12, Jul 2019.
- [4] Jonathan Masci, Davide Boscaini, Michael M. Bronstein, and Pierre Vandergheynst. Geodesic convolutional neural networks on riemannian manifolds, 2015.
- [5] D. Maturana and S. Scherer. Voxnet: A 3d convolutional neural network for real-time object recognition. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 922–928, Sep. 2015.

- [6] Hang Su, Subhransu Maji, Evangelos Kalogerakis, and Erik G. Learned-Miller. Multi-view convolutional neural networks for 3d shape recognition. In *Proc. ICCV*, 2015.
- [7] Petar Veličković, William Fedus, William L. Hamilton, Pietro Liò, Yoshua Bengio, and R Devon Hjelm. Deep graph infomax, 2018.
- [8] Peng-Shuai Wang, Yang Liu, Yu-Xiao Guo, Chun-Yu Sun, and Xin Tong. Adaptive O-CNN: A Patch-based Deep Representation of 3D Shapes. *ACM Transactions on Graphics (SIGGRAPH Asia)*, 37(6), 2018.