3D Point Cloud Analysis

Shan Liu • Min Zhang • Pranav Kadam
C.-C. Jay Kuo

3D Point Cloud Analysis

Traditional, Deep Learning, and Explainable Machine Learning Methods

Springer

Shan Liu
Tencent Media Lab
Palo Alto, CA, USA

Min Zhang
University of Southern California
Los Angeles, CA, USA

Pranav Kadam
University of Southern California
Los Angeles, CA, USA

C.-C. Jay Kuo
University of Southern California
Los Angeles, CA, USA

ISBN 978-3-030-89179-4 ISBN 978-3-030-89180-0 (eBook)
https://doi.org/10.1007/978-3-030-89180-0

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Dedicated to my son, William

—Shan Liu

Dedicated to my parents and grandparents for their love and support

—Min Zhang

Dedicated to my parents for their love, encouragement, and support

—Pranav Kadam

Dedicated to my wife, Terri, and my daughter, Allison

—C.-C. Jay Kuo

Preface

Three-dimensional (3D) point clouds are gaining increasing attention for the emerging applications of 3D vision. Point clouds have widespread use in several spectrums of fields, include robotics, 3D graphics, autonomous driving, virtual reality, and so on. To keep pace with the increasing applications, the research and development of methods and algorithms to effectively store, process, and infer meaning from point cloud is on the rise. The traditional algorithms for analyzing point clouds focus on encoding the local geometric properties of points. The success of deep learning methods for processing image data led to similar networks being developed for point clouds. Present day research heavily involves the development of deep networks for various point cloud processing tasks.

The aim of this book is to give a high-level overview of point clouds and acquaint the reader with some of the most popular methods and techniques for point cloud processing. The ideal audience are those with a basic knowledge of linear algebra, machine learning, and deep learning algorithms, who wish to explore point clouds in their career or as a hobby.

This book is organized into five chapters. Chapter 1 introduces 3D point clouds and various related tasks. Chapter 2 discusses traditional point cloud analysis, including some basic operations such as filtering, nearest neighbor searching, and model fitting techniques, along with feature detectors and descriptors. Chapter 3 on deep learning discusses some of the most common machine learning-based methods. The deep learning literature is abundant, with more research being published as we write this book. We discuss some of the most representative methods that summarize the overall research direction. The emphasis is on understanding the model architecture and the novelty. The experimental details are omitted, and only key results from papers are provided. Chapter 4 on explainable machine learning methods presents our own research, which is based on a new machine learning paradigm called successive subspace learning (SSL). SSL offers several advantages over deep learning methods. Enough background review on SSL is provided prior

to a thorough discussion of SSL-based methods for point cloud processing. Some applications of SSL to other vision tasks are also discussed. The final chapter (Chap. 5) includes a summary and some concluding remarks as well as possible future research directions.

Contents

1 Introduction 1
1.1 Introduction 1
1.2 3D Point Clouds 2
1.2.1 Point Cloud Formation 2
1.2.2 Comparison with Other Visual Data Forms 3
1.3 Point Cloud Processing 4
1.3.1 Registration 4
1.3.2 Classification 6
1.3.3 Semantic Segmentation 6
1.3.4 Odometry 6
1.4 Applications 7
1.5 Datasets 8
1.5.1 ModelNet40 8
1.5.2 ShapeNet 9
1.5.3 S3DIS 9
1.5.4 3D Match 10
1.5.5 KITTI 10
1.6 Summary 12
References 12

2 Traditional Point Cloud Analysis 15
2.1 Filtering 15
2.1.1 Downsampling 15
2.1.2 Noise Removal 18
2.2 Nearest Neighbor Search 21
2.2.1 Binary Search Tree 22
2.2.2 k-Dimensional Tree 23
2.2.3 Octree 26
2.3 Model Fitting 28
2.3.1 Least Squares Fitting 28
2.3.2 Hough Transform 31
2.3.3 Random Sample Consensus 33

2.4 Point Cloud Features 36
2.4.1 Feature Detectors 36
2.4.2 Feature Descriptors 42
2.5 Classification and Segmentation 45
2.6 Registration 46
2.6.1 Iterative Closest Point (ICP) 47
2.6.2 Point-to-Plane ICP 48
2.6.3 Generalized ICP 49
2.6.4 Global Registration 50
References 51

3 Deep Learning-Based Point Cloud Analysis 53
3.1 Introduction 53
3.2 Classification and Segmentation 54
3.2.1 PointNet 54
3.2.2 PointNet++ 56
3.2.3 Dynamic Graph CNN 58
3.2.4 PointCNN 60
3.2.5 PointSIFT 62
3.2.6 Point Transformer 64
3.2.7 RandLA-Net 66
3.3 Registration 69
3.3.1 PointNetLK 70
3.3.2 Deep Closest Point 72
3.3.3 PRNet 75
3.3.4 3D Match 78
3.3.5 PPFNet 79
3.3.6 Deep Global Registration 81
References 84

4 Explainable Machine Learning Methods for Point Cloud Analysis 87
4.1 Successive Subspace Learning on 2D Images 87
4.1.1 Data-Driven Saak Transform 88
4.1.2 Handwritten Digit Recognition by Saak Transform 89
4.1.3 Interpretable Convolutional Neural Networks via Feedforward Design 90
4.1.4 PixelHop 92
4.1.5 PixelHop++ 93
4.2 Classification and Part Segmentation 96
4.2.1 PointHop 97
4.2.2 PointHop++ 103
4.2.3 Unsupervised Feedforward Feature (UFF) 108
4.3 Registration 113
4.3.1 Salient Points Analysis (SPA) 115
4.3.2 R-PointHop 121

4.4 Other Applications of Successive Subspace Learning 135
4.4.1 FaceHop 135
4.4.2 DefakeHop 136
4.4.3 AnomalyHop 137
References 138

5 Conclusion and Future Work 141
5.1 Conclusion 141
5.2 Future Work 142

Index 145

Author Biographies

Shan Liu received the B.Eng. degree in Electronic Engineering from the Tsinghua University and the M.S. and Ph.D. degrees in Electrical Engineering from the University of Southern California. She is currently a Distinguished Scientist at the Tencent and General Manager of the Tencent Media Lab. She was the former Director of Media Technology Division at MediaTek, USA. She was also formerly with the MERL and Sony, etc. Dr. Liu has been actively contributing to international standards for more than a decade. She has numerous technical proposals adopted into various standards, such as H.266/VVC, H.265/HEVC, OMAF, DASH, MMT, and PCC, and served as co-editor of H.265/HEVC SCC and H.266/VVC. Meanwhile, technologies and products developed by her and her team have served hundreds of millions of users. Dr. Liu holds more than 200 granted patents and has published more than 100 technical papers. She was named "APSIPA Industrial Distinguished Leader" by the Asia-Pacific Signal and Information Processing Association in 2018, and "50 Women in Tech" by the Forbes China in 2020. She is on the Editorial Board of IEEE Transactions on Circuits and Systems for Video Technology (2018–present) and received the Best AE Award in 2019 and 2020. Her research interests include audio-visual, volumetric, immersive, and emerging media compression, intelligence, transport, and systems.

Min Zhang received her B.E. degree from the School of Science, Nanjing University of Science and Technology, Nanjing, China, and her M.S. degree from the Viterbi School of Engineering, University of Southern California (USC), Los Angeles, USA, in 2017 and 2019, respectively. She joined the Media Communications Laboratory (MCL) in the summer of 2018 and is currently a Ph.D. student in the USC, guided by Prof. C.-C. Jay Kuo. Her research interests include point cloud processing and analysis-related problems, such as point cloud classification, registration, and segmentation and detection, in the field of 3D computer vision, machine learning, and perception.

Pranav Kadam received his M.S. degree in Electrical Engineering from the University of Southern California, Los Angeles, USA, in 2020, and the Bachelor's degree in Electronics and Telecommunication Engineering from the Savitribai Phule Pune University, Pune, India, in 2018. He is currently pursuing the Ph.D. degree in Electrical Engineering from the University of Southern California. He is actively involved in the research and development of methods for point cloud analysis and processing. His research interests include 3D computer vision, machine learning, and perception.

C.-C. Jay Kuo received the Ph.D. degree in Electrical Engineering from the Massachusetts Institute of Technology, Cambridge, in 1987. He is currently the holder of William M. Hogue Professorship, a Distinguished Professor of Electrical and Computer Engineering and Computer Science, and the Director of the USC Multimedia Communications Laboratory (MCL) at the University of Southern California. Dr. Kuo is a Fellow of the American Association for the Advancement of Science (AAAS), the Institute of Electrical and Electronics Engineers (IEEE), the National Academy of Inventors (NAI), and the International Society for Optical Engineers (SPIE). He has received several awards for his research contributions, including the 2010 Electronic Imaging Scientist of the Year Award, the 2010–2011 Fulbright-Nokia Distinguished Chair in Information and Communications Technologies, the 2011 Pan Wen-Yuan Outstanding Research Award, the 2019 IEEE Computer Society Edward J. McCluskey Technical Achievement Award, the 2019 IEEE Signal Processing Society Claude Shannon-Harry Nyquist Technical Achievement Award, the 2020 IEEE TCMC Impact Award, the 72nd annual Technology and Engineering Emmy Award (2020), and the 2021 IEEE Circuits and Systems Society Charles A. Desoer Technical Achievement Award.