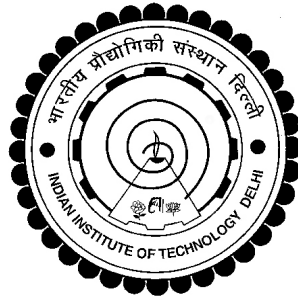


FEATURES FOR 3D POINT CLOUDS

SIDDHARTH SRIVASTAVA



**DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY DELHI
FEBRUARY 2019**

©Indian Institute of Technology Delhi (IITD), New Delhi, 2019

FEATURES FOR 3D POINT CLOUDS

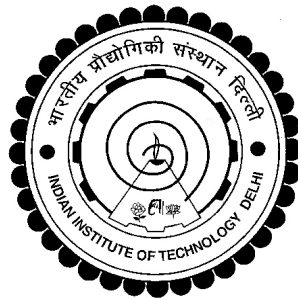
by

SIDDHARTH SRIVASTAVA

DEPARTMENT OF ELECTRICAL ENGINEERING

Submitted

*in fulfillment of the requirements of the degree of Doctor of Philosophy
to the*



INDIAN INSTITUTE OF TECHNOLOGY DELHI

FEBRUARY 2019

Certificate

This is to certify that the thesis entitled “**Features for 3D Point Clouds**” being submitted by **Mr. Siddharth Srivastava** to the Department of Electrical Engineering, Indian Institute of Technology Delhi, for the award of the degree of **Doctor of Philosophy** is the record of the bona-fide research work carried out by him under my supervision. In my opinion, the thesis has reached the standards fulfilling the requirements of the regulations relating to the degree.

Prof. Brijesh Lal

Department of Electrical Engineering

Indian Institute of Technology Delhi

Acknowledgements

I would like to thank my supervisor Prof. Brejesh Lall for his continuous support during the period of my stay at IIT Delhi. He has not only provided guidance in technical progress, but has also broadened the spectrum of my thinking. He has patiently listened to numerous of my ideas, and with his ability to transform ideas into potential research problems, he has transformed me from being a eager student to a competent researcher Prof. Lall's calm demeanour and exceptional ability to understand his students, has been a life changing experience for me.

I would also like to thank my friends Sumit, Prerana, Swati, Parul, Rahul, Anurag and my team at C-DAC for their continuous support and listening to endless monologues on my research problems. Towards the last, I would like to mention that this dissertation would not have been possible without support of my parents and sister, who have continuously kept me motivated and have made numerous sacrifices during this period.

Siddharth Srivastava

Abstract

In this dissertation, we develop compact, robust and efficient local descriptors to represent 3D models captured from devices such as Kinect, LiDAR *etc.* Among many possible input data representations for 3D data, we choose 3D point clouds as the preferred form of data representation for the work in this dissertation. 3D point clouds have been chosen as they are the raw output of many 3D scanning devices and hence provide the closest representation of the captured scenes or objects. Our first contribution towards developing compact and robust local descriptor is developing a technique for extracting a binary descriptor directly from 3D point clouds. For construction of the descriptor, we introduce an angular criterion for finding the points representing the geometry around a keypoint. The information from these points is encoded in a pairwise manner by comparing the projection of normals. As a second contribution, we propose a more robust and fast to compute binary descriptor. It involves partitioning the space around keypoints into sectors and performing hierarchical pairwise comparison of projection of normals. Additionally, to reduce the false positives we propose fitting a B-Spline surface to the neighbourhood of keypoints and matching keypoints with similar B-Spline surfaces. The similar surfaces are obtained by comparing residual error between possible cubic surfaces and constructing a look-up table in an offline manner. While with handcrafted descriptors we were able to encode the geometry, we observed that the progress in deep networks for tasks such as 3D object classification began providing better than state-of-the-art results. However, such methods do not compare their performance against techniques which involve handcrafted descriptors, and also do not always provide results on standard benchmarks. To fill this gap, the third contribution is a comprehensive evaluation of such networks on common benchmarks and studying the impact of binarization on the descriptors obtained from such networks. Our study indicates that the descriptors constructed using deep networks provide strong feature representation. However, these descriptors are global in nature and further mostly work on voxels or projections obtained from 3D point clouds. Moreover, such networks strongly rely on availability of large amount of data. On the contrary, handcrafted descriptors may even be used for unconventional tasks by adjusting parameters. Therefore, to analyze the applicability of deep networks to such tasks, we apply them for drought stress classification using 3D plant models. As plants usually have very sharp structural variations, this task also puts to test the capability of deep networks to discriminately represent

complex object structures. The experiments indicate that deep networks perform better when local information is infused within the descriptors derived from them. Motivated by these experiments, as a fourth contribution we develop a local descriptor based on deep networks and propose a multi-margin loss function to exploit the complexity in training data. The proposed descriptor is evaluated for various tasks on standard benchmarks against competing methods. In summary, this dissertation proposes state-of-the-art handcrafted and deep network based 3D point cloud descriptors. The dissertation also contains a comparison of various methods against standard baselines.

सार

इस शोध प्रबंध में, हम 3डी डेटा के लिए कई संभावित इनपुट डेटा अभ्यावेदन जैसे किक्टेक्ट, लीडर आदि से कैप्चर किए गए 3डी मॉडल का प्रतिनिधित्व करने के लिए कॉम्पैक्ट, मजबूत और कुशल स्थानीय डिस्क्रीप्टर विकसित करते हैं, हम डेटा प्रतिनिधित्व के लिए पसंदीदा फॉर्म के रूप में 3डी पॉइंट क्लाउड चुनते हैं इस शोध प्रबंध में काम करते हैं। 3डी बिंदु बादलों को चुना गया है क्योंकि वे कई 3डी स्कैनिंग उपकरणों के कच्चे आउटपुट हैं और इसलिए कैप्चर किए गए दृश्यों या वस्तुओं का निकटतम प्रतिनिधित्व प्रदान करते हैं।

कॉम्पैक्ट और मजबूत स्थानीय डिस्क्रीप्टर विकसित करने की दिशा में हमारा पहला योगदान 3डी पॉइंट क्लाउड से सीधे बाइनरी डिस्क्रीप्टर निकालने की तकनीक विकसित करना है। डिस्क्रीप्टर के निर्माण के लिए, हम एक कीप के चारों ओर ज्यामिति का प्रतिनिधित्व करने वाले बिंदुओं को खोजने के लिए कोणीय मानदंड का परिचय देते हैं। इन बिंदुओं से जानकारी मानदंडों के प्रक्षेपण की तुलना करके एक जोड़ीदार तरीके से एन्कोडेड है। दूसरे योगदान के रूप में, हम बाइनरी डिस्क्रीप्टर की गणना करने के लिए एक अधिक मजबूत और तेजी का प्रस्ताव देते हैं। इसमें मुख्य बिंदुओं के आसपास के क्षेत्रों को सेक्टरों में विभाजित करना और मानदंडों के प्रक्षेपण के पदानुक्रमित जोड़ीवार प्रदर्शन करना शामिल है। इसके अतिरिक्त, झूठी सकारात्मकता को कम करने के लिए हम एक बी-स्पाइन सतह को फिटिंग करने का प्रस्ताव करते हैं जो कि के-पॉइंट्स के पड़ोस में है और समान बी-स्प्लिन सतहों के साथ कीपॉइंट्स का मिलान करता है। समान सतहों को संभव घन सतहों के बीच अवशिष्ट त्रुटि की तुलना करके और ऑफ़लाइन तरीके से एक लुक-अप तालिका का निर्माण करके प्राप्त किया जाता है।

जबकि दस्तकारी वर्णनकर्ताओं के साथ हम ज्यामिति को एन्कोड करने में सक्षम थे, हमने देखा कि 3डी ऑब्जेक्ट वर्गीकरण जैसे कार्यों के लिए गहरे नेटवर्क में प्रगति अत्याधुनिक परिणामों से बेहतर प्रदान करना शुरू कर दिया। हालांकि, ऐसे तरीके तकनीक के खिलाफ उनके प्रदर्शन की तुलना नहीं करते हैं जिसमें दस्तकारी वर्णनकर्ता शामिल होते हैं, और हमेशा मानक बेंचमार्क पर परिणाम प्रदान नहीं करते हैं। इस अंतर को भरने के लिए, तीसरा योगदान आम बेंचमार्क पर इस तरह के नेटवर्क का एक व्यापक मूल्यांकन है और ऐसे नेटवर्क से प्राप्त विवरणकर्ताओं पर द्वैरीकरण के प्रभाव का अध्ययन करना है। हमारे अध्ययन से संकेत मिलता है कि गहरे नेटवर्क का उपयोग करके निर्मित वर्णनकर्ता मजबूत सुविधा प्रतिनिधित्व प्रदान करते हैं। हालांकि, ये वर्णनकर्ता प्रकृति में वैश्विक हैं और आगे अधिकतर डी बिंदु बादलों से प्राप्त स्वर या अनुमानों पर काम करते हैं। इसके अलावा, इस तरह के नेटवर्क दृढ़ता से बड़ी मात्रा में डेटा की उपलब्धता पर निर्भर करते हैं। इसके विपरीत, मापदंडों को समायोजित करके अप्रकाशित कार्यों के लिए दस्तकारी वर्णनकर्ताओं का उपयोग भी किया जा सकता है। इसलिए, ऐसे कार्यों के लिए गहरे नेटवर्क की प्रयोज्यता का विश्लेषण करने के लिए, हम उन्हें 3डी प्लॉट मॉडल का उपयोग करके सूखे के तनाव वर्गीकरण के लिए लागू करते हैं। जैसा कि पौधों में आमतौर पर बहुत तेज संरचनात्मक विविधताएं होती हैं, यह कार्य जटिल वस्तु संरचनाओं का भेदभाव करने के लिए गहरे नेटवर्क की क्षमता का परीक्षण करने के लिए भी डालता है। प्रयोगों से संकेत मिलता है कि गहरे नेटवर्क बेहतर प्रदर्शन करते हैं जब स्थानीय जानकारी उनसे प्राप्त विवरणों के भीतर डाली जाती है। इन प्रयोगों से प्रेरित होकर, चौथे योगदान के रूप में हम गहरे नेटवर्क पर आधारित एक स्थानीय डिस्क्रीप्टर विकसित करते हैं और प्रशिक्षण डेटा में जटिलता का फायदा उठाने के लिए एक बहु-मार्जिन हानि फ़ंक्शन का प्रस्ताव करते हैं। प्रतियोगी विधियों के खिलाफ मानक बेंचमार्क पर विभिन्न कार्यों के लिए प्रस्तावित विवरणक का मूल्यांकन किया जाता है। सारांश में, यह शोध प्रबंध अत्याधुनिक दस्तकारी और डीप नेटवर्क आधारित 3डी पॉइंट क्लाउड डिस्क्रीप्टर का प्रस्ताव करता है। शोध प्रबंध में मानक आधारभूत के खिलाफ विभिन्न तरीकों की तुलना भी शामिल है।

Contents

Certificate	i
Acknowledgements	ii
Abstract	iii
List of Figures	x
List of Tables	xii
1 Introduction	1
2 Literature Survey	7
2.1 Introduction	7
2.2 Keypoint Detection	9
2.3 Handcrafted Descriptors	11
2.4 Deep Learning based Global 3D Descriptors	13
2.5 Deep Learning based Local 3D Descriptors	14
2.6 Conclusion	14
3 3D Binary Signatures	16
3.1 Introduction	16
3.2 Methodology	17
3.2.1 Keypoint Detection	18
3.2.2 Nearest Neighbours with Angular Constraint	19
3.2.3 Alignment with Local Reference Frame	20
3.2.4 Descriptor Generation	20

3.2.5	Descriptor Matching	21
3.3	Experiments and Results	23
3.3.1	Experimental Setup	23
3.3.2	Performance Evaluation	23
3.4	Conclusion	27
4	Extending 3D Binary Signatures	29
4.1	Introduction	29
4.2	Models	30
4.2.1	Extended 3D Binary Signatures (E3DBS)	30
4.2.2	B-Spline surface based matching of 3D Binary Signatures (S3DBS)	33
4.3	Experiments and Results	36
4.3.1	Parameter and Implementation Details	36
4.3.2	Performance Evaluation	37
4.4	Conclusion	39
5	DeepPoint3D: Learning Discriminative Local Descriptors using Deep Metric Learning on 3D Point Clouds	41
5.1	Introduction	41
5.2	Methodology	42
5.2.1	Overview	42
5.2.2	Network Input	43
5.2.3	Loss Function	44
5.2.4	Proposed Networks	46
5.2.5	Network Training	46
5.3	Experiments	47
5.3.1	Dataset	47
5.3.2	Evaluation Metrics	48
5.3.3	Parameter Settings and Evaluation Setup	49
5.3.4	Discussion	49
5.4	Applications	51
5.4.1	Classification	54
5.4.2	Retrieval	54

5.4.3	Binarization	55
5.4.4	Drought Stress Classification	56
5.5	Conclusion	56
6	Empirical Evaluation of Recent Learning based Methods for 3D Shape Retrieval and Classification	58
6.1	Introduction	58
6.2	Evaluated Methods	60
6.3	Experimental Evaluations	64
6.3.1	Experimental Setup	64
6.3.2	Classification on ModelNet10 and ModelNet40	65
6.3.3	Retrieval on ModelNet10 and ModelNet40	68
6.3.4	Discussion	70
6.4	Conclusion	73
7	Drought Stress Classification using 3D Plant Models	75
7.1	Introduction	75
7.2	Methodology	76
7.2.1	3D Plant Reconstruction using Learned Invariant Feature Transform	76
7.2.2	3D Segmentation using Voxel Cloud Connectivity Segmentation	78
7.2.3	3D Features: Local and Deep Descriptors	79
7.2.4	Drought Stress Classification	81
7.3	Experimental Results	82
7.3.1	Dataset	82
7.3.2	Results	82
7.4	Conclusion	85
8	Conclusions and Future Work	86
	Bibliography	88
	List of Publications	99

List of Figures

2.1	Example showing utilization of descriptors for object classification task	8
2.2	Pipeline of Fixed Scale Keypoint Detectors	9
2.3	Pipeline of Scale Invariant Keypoint Detectors	10
3.1	Construction of 3D Binary Signatures	18
3.2	Visualization of a) Angular Constraint b) Comparison of projection of Normals	19
3.3	Compactness of Descriptors	26
3.4	Matching time comparison	27
3.5	Performance on increasing NN	28
4.1	Space subdivision for constructing E3DBS	31
4.2	B-Spline surface based matching of descriptors	34
4.3	Compactness of Descriptors	38
4.4	Computation Time (sec) of Descriptors	39
5.1	Proposed deep architectures for generating local descriptors directly from 3D Point Clouds	43
5.2	Correspondence Accuracy on varying dimension of the descriptor with multi-margin contrastive loss.	51
6.1	Impact on classification on varying k in kNN algorithm. The results are shown for 64bit ITQ binary vector on ModelNet40.	73
7.1	Flow diagram for the proposed Drought Stres Classification approach.	76
7.2	Sample views of a wheat plant under control and drought stress condition in the phenomics facility.	80

List of Tables

3.1	Datasets used in the study	23
3.2	Area under the Precision-recall curve (AUC_{PR})	25
3.3	Descriptor Length	26
4.1	Area under the Precision-Recall curve (AUC_{PR})	38
4.2	Area under the Precision-Recall curve with B-Spline surface matching (AUC_{PR})	38
4.3	Descriptor Length	39
5.1	Performance of the proposed architectures on the test dataset.	51
5.2	Comparison with state of the art at default mesh resolution.	52
5.3	CMC and Correspondence Accuracy on test dataset at $mr = 0.25$ for deep network based methods.	52
5.4	CMC and Correspondence Accuracy on test dataset at $mr=0.5$ for deep network based methods.	53
5.5	CMC and Correspondence Accuracy on test Dataset with $mr = 0.75$ for deep network based methods.	53
5.6	Classification Accuracy on Semantic3D with local descriptors	54
5.7	Retrieval mAP on ModelNet10	55
5.8	Classification and Retrieval Performance on binary descriptor.	56
5.9	Classification Accuracy in 3D and Feature Computation Time	57
6.1	Comparison of State of the art for Classification Accuracy(%) on Mod- elNet10 and ModelNet40.	61
6.2	Comparison of State of the art Retrieval Accuracy (mAP) on Model- Net10 and ModelNet40.	61

6.3	Classification accuracy of descriptors from deep networks on quantization	70
6.4	Performance of binarized feature vectors for Retrieval (mAP) on ModelNet10 and ModelNet40.	71
6.5	Retrieval performance (mAP %) on ModelNet10 and ModelNet40 datasets.	71
6.6	Comparison on ModelNet40 of average retrieval time per model for real valued vs 128 bit binary vector (ITQ) for various methods	72
7.1	Classification Accuracy in 3D and Feature Computation Time	83