REPORT DOCUMENTATION PAGE	Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
05-03-2017	Final Report	12-Aug-2013 - 11-Dec-2016

4. TITLE AND SUBTITLE	
Final Report: Research area:3.1 Modeling of Complex Systems, 3.1.1 Geometric and Topological Modeling; Title: 3D Modeling of Urban Sites from Point Clouds	5a. CONTRACT NUMBER W911NF-13-1-0358
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER 611102
6. AUTHORS Suya You	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES	8. PERFORMING ORGANIZATION REPORT NUMBER
University of Southern California Contracts & Grants 3720 S. Flower St. Los Angeles, CA 90089 -0701	

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211	ARO
	11. SPONSOR/MONITOR'S REPORT NUMBER(S) 63604-CS.9

12. DISTRIBUTION AVAILIBILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT

This project focuses on the developments of advanced techniques for processing 3D point cloud data specially targeted for 3D urban modeling. Current work on urban modeling has focused on either isolated objects or large-scale structures, not the wide diversity of small-scale objects occurring in urban scenes. The ultimate objective of this research is to pursue new techniques and solutions that extend the range of object classes that can be rapidly discovered, recognized, extracted and modeled from point-cloud data.

15. SUBJECT TERMS

3D modeling, object detection, object recognition, shape matching, deep learning, point clouds

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Suya You
UU	UU	UU	UU		19b. TELEPHONE NUMBER 213-740-4495

Standard Form 298 (Rev 8/98)
Prescribed by ANSI Std. Z39.18

Report Title

Final Report: Research area:3.1 Modeling of Complex Systems, 3.1.1 Geometric and Topological Modeling; Title: 3D Modeling of Urban Sites from Point Clouds

ABSTRACT

This project focuses on the developments of advanced techniques for processing 3D point cloud data specially targeted for 3D urban modeling. Current work on urban modeling has focused on either isolated objects or large-scale structures, not the wide diversity of small-scale objects occurring in urban scenes. The ultimate objective of this research is to pursue new techniques and solutions that extend the range of object classes that can be rapidly discovered, recognized, extracted and modeled from point-cloud data.

Enter List of papers submitted or published that acknowledge ARO support from the start of the project to the date of this printing. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

Received	Paper

TOTAL:

Number of Papers published in peer-reviewed journals:

(b) Papers published in non-peer-reviewed journals (N/A for none)

Received	Paper

TOTAL:

Number of Papers published in non peer-reviewed journals:

(c) Presentations

Number of Presentations: 0.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received		Paper
03/05/2017	7	Jing Huang and Suya You. Vehicle Detection in Urban Point Clouds with Orthogonal- View Convolutional Neural Network, IEEE International Conference on Image Processing (ICIP). 25-SEP-16, Phoenix, AZ. : ,
03/05/2017	8	Jing Huang and Suya You. Point Cloud Labeling using 3D Convolutional Neural Network, International Conference on Pattern Recognition (ICPR). 04-DEC-16, Cancun, Mexico. : ,

TOTAL: **2**

Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received		Paper
08/29/2014	1.00	Jing_Huang, Suya_You. Detecting Objects in Scene Point Cloud: A Combinational Approach, International Conference on 3D Vision (3DV). 29-JUN-13, . : ,
08/29/2014	2.00	Wei_Guan, Suya_You. Robust Image Matching with Line Context, The 24th British Machine Vision Conference (BMVC 2013). 09-SEP-13, . : ,
08/29/2014	3.00	Jing_Huang, Suya_You. Segmentation and Matching: Towards a Robust Object Detection System, IEEE Winter Conference on Applications of Computer Vision (WACV). 24-MAR-14, . : ,
10/09/2015	4.00	Suya You, Jing Huang. Detection and Recognition of Pole-Like Urban Objects from Point Clouds, IEEE International Conference on Robotics and Automation (ICRA). 26-MAY-15, . : ,
10/09/2015	5.00	Jing Huang,, Suya You. Change Detection in Laser-Scanned Data of Industrial Sites, IEEE Winter Conference on Applications of Computer Vision (WACV). 06-JAN-15, . : ,
10/09/2015	6.00	Guan Pang, , Rongqi Qiu, , Jing Huang, , Suya You, , Ulrich Neumann. Automatic 3D Industrial Point Cloud Modeling and Recognition, IAPR InternationalConference on Machine Vision Applications (ICMVA). 18-MAY-15, . : ,

TOTAL: **6**

Number of Peer-Reviewed Conference Proceeding publications (other than abstracts):

(d) Manuscripts

<u>Received</u> <u>Paper</u>

TOTAL:

Number of Manuscripts:

Books

<u>Received</u> <u>Book</u>

TOTAL:

<u>Received</u> <u>Book Chapter</u>

TOTAL:

Patents Submitted

1. Guan Pang, Jing Huang, Suya You, and Ulrich Neumann, “Multi-view 3D Object Recognition From a Point Cloud and Change Detection”, US patent application, filed on March 2015

2. Suya You, Jing Huang, Amir Anvar, and Christopher Fisher, “System and Method of Detecting Objects in Scene Point Cloud”, US patent application, filed on 2013

3. Ulrich Neumann, Suya You, “Three-Dimensional Point Processing and Model Generation”, U.S. Patent Application, filled on 2013

Patents Awarded

Awards

Graduate Students

NAME	PERCENT_SUPPORTED	Discipline
Jing Huang	0.50	
Weiyue Wang	0.50	
FTE Equivalent:	**1.00**	
Total Number:	**2**	

Names of Post Doctorates

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Names of Faculty Supported

NAME	PERCENT_SUPPORTED	National Academy Member
Suya You	0.27	
Ulrich Neumann	0.10	
FTE Equivalent:	**0.37**	
Total Number:	**2**	

Names of Under Graduate students supported

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Student Metrics

This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period: 0.00

The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:...... 0.00

The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:...... 0.00

Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):...... 0.00

Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:...... 0.00

The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense 0.00

The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields:...... 0.00

Names of Personnel receiving masters degrees

NAME
Xiang Li
Total Number: 1

Names of personnel receiving PHDs

NAME
Jing Huang
Total Number: 1

Names of other research staff

NAME	PERCENT_SUPPORTED

FTE Equivalent:
Total Number:

Sub Contractors (DD882)

Inventions (DD882)

Scientific Progress

See Attachment

Technology Transfer

See Attachment

Final Report

(08/2013 – 08/2016)

3D Modeling of Urban Sites from Point Clouds

Suay You, Ulruch Neumann

University of Southern California

1. Summary

This project focuses on the developments of advanced techniques and methods for processing 3D point-cloud data specially targeted for 3D urban modeling. Current work on urban modeling has focused on either isolated objects or large-scale structures such as buildings and roads, not the wide diversity of small-scale objects occurring in urban scenes. An important capability to military missions is rapid awareness and modeling of a multitude of commonly occurring objects in urban environments such as lamp posts, street lights, traffic signs, and vehicles. The ultimate objective of this research is to pursue new techniques and solutions that extend the range of urban object classes that can be rapidly discovered, recognized, extracted and modeled from point-cloud data.

The pursued technical approach is a novel alternative to traditional modeling approaches. The novelty arises from using the strategy of Modeling by Recognition (MBR) to rapidly identify objects from a 3D library of objects within point-cloud data. The recognized-object point clouds are then replaced with library data, such as polygon surface models, thereby constructing accurate and complete 3D scene models. The MBR technique has significant advances over traditional modeling methods in terms of modeling accuracy, efficiency, and scalability. The emphasis is to provide not only the capability to rapidly create the 3D urban models as realistically as possible, but also to provide semantic labels of objects and the ability to update the model library.

The research foci are the key components in the proposed modeling approach: (1) MBR framework to rapidly process point-cloud data and construct accurate and complete 3D models, and (2) robust 3D shape matching algorithms that are used to detect objects of interest from point-cloud inputs and match them to model-library elements.

The research efforts cover all the proposed tasks that include core algorithm developments, MBR system integration, and testing and evaluations. We extended the identifiable classes of urban objects of interest and developed techniques for detecting and modeling these urban objects. We studied the mechanism and properties of image self-similarity and developed the theorical interpretations for image self-similarity in features representation and shape matching. We developed the novel approaches of pole-like object detection, vehicle detection, semantic labeling, and change detection, and integrated them into MBR system and evaluated with various datasets. We completed all the proposed tasks. We produced 8 peer-reviewed publications and 3 patent applications resulted from this research effort. We presented and demonstrated the developed new work to several defense contractors and industries for technology transfer and collaborations. In addition, one Ph.D. student supported by this project successfully passed his Ph.D. defense and three MS students who are partially supported by this project graduated.

2. Description of Research Efforts and Achievements

2.1. 3D shape representation and description

We developed a new method for 3D shape description and matching from point clouds. The new method (named 3D-SSIM) is a novel local feature representation and matching process based on the concept of shape *self-similarities*. Self-similarity is a unique property of fractals and topological geometry. It captures the internal geometric layout of local patterns. Locations in images with self-similarity form local patterns distinguishable from neighbor locations, which can greatly facilitate matching across images that appear substantially different in visual appearance. We have evaluated extensively the 3D-SSIM method with various point-cloud datasets, and compared with state-of-the-art methods (i.e. Spin Image, Shape Context, Point Feature Histograms, 3D SFIT). The results show that the 3D self-similarity method leads to a robust feature representation that is both geometry and illumination invariant by modeling, reproducing and compensating for substantial changes between point cloud datasets. Figure 1 and Figure 2 show the results with simulation data and real-world data.

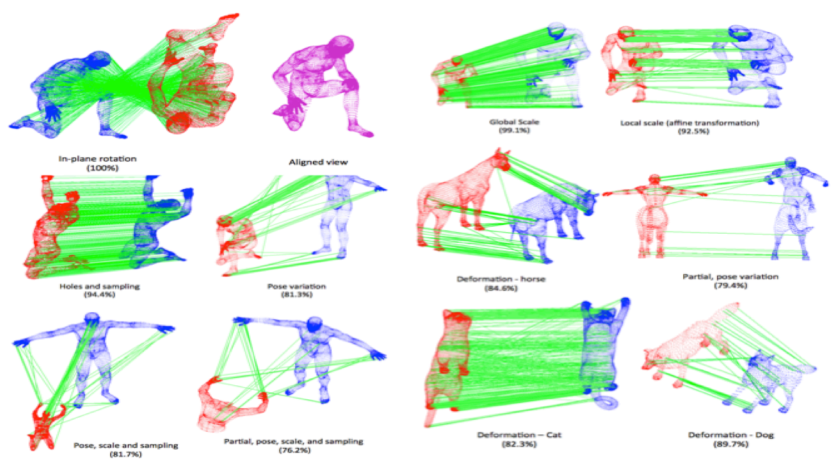

Figure 1: Quantitative evaluation results of 3D-SSIM with SHREC benchmark data

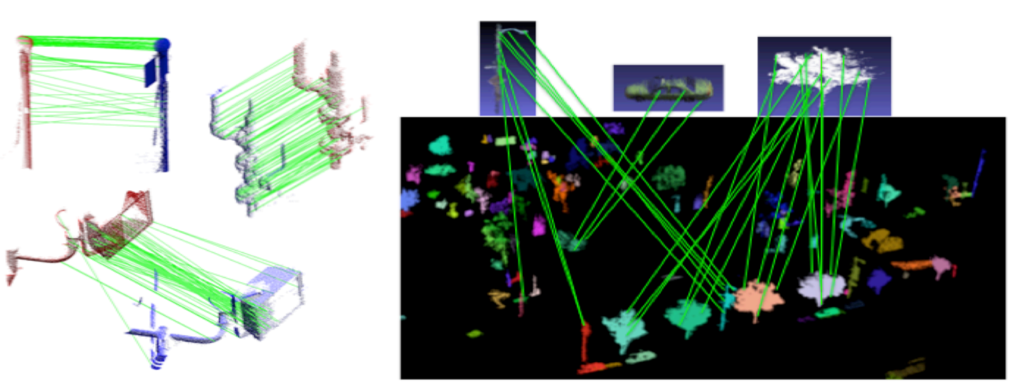

Figure 2: Apply 3D-SSIM to match and recognize various urban objects

2.2. Robust 3D shape matching and recognition

We developed an object detection and recognition technique combining learning-based classification and 3D-SSIM local features for detecting multiple objects in complex point clouds. We explore the complementary properties of these two techniques to develop a high-level detection and recognition strategy.

We observe that a majority of urban man-made objects in existence are composed of primitives that often have planarity and regularity properties. This majority of object should be detected first to form a simple and clean representation of these scene elements. For example, in aerial LiDAR of urban scenes, roofs and ground are often piecewise planar. Similarly, in street-level LiDAR, facades and ground are often planar. Use of local feature matching for planar surfaces is unlikely to succeed due to the lack of discriminating local features. Instead, our strategy is to extract the primitive shapes using their global and high-level geometric properties. We developed a hybrid approach combining the learn-based SVM classification and 3D-SSIM local feature-based method to first infer the major primitive shapes, and then use the 3D-SSIM local feature matching to detect the residual objects using the constraints provided from the primitive shapes for the search for associated objects and propagation of spatial relationships. Figure 3 shows the performance evaluation results on Princeton benchmark dataset, which indicates our approach is superior to the state-of-the-art methods including Spin Image (SI), Shape Context (SC), and Fast Point Feature Histograms (FPFH).

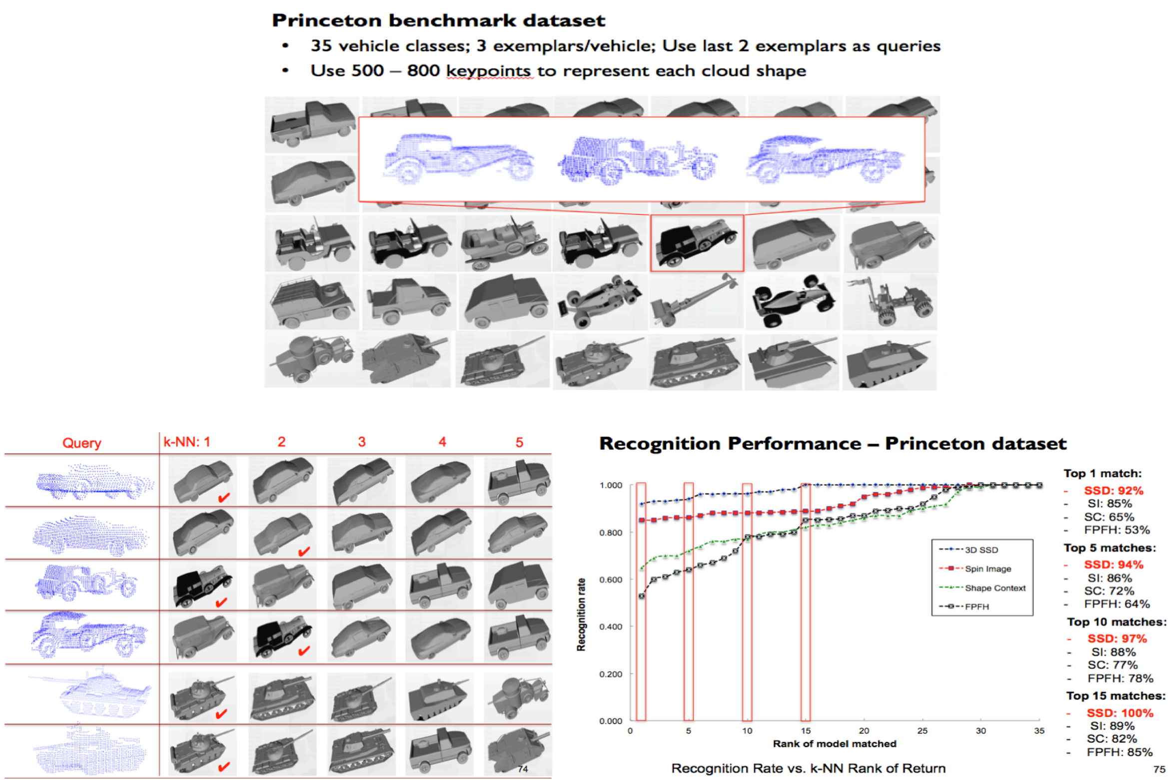

Figure 3: Evaluation of hybrid object recognition using Princeton benchmark dataset

2.3. Detection and modeling of pole-like urban objects

We identified classes of commonly occurring objects in urban sites and classified the objects according to their geometries and functionalities as three classes: *pole-like objects* (e.g. street lamps, traffic lights, traffic signs, posts, etc.), *box-like objects* (e.g. mailbox, trashcans, fire hydrants, etc.), and *dynamic objects* (e.g. vehicles, people, etc.). The object classes are updated and grown constantly to form a complete hierarchy of shapes and objects in complex urban sites.

We developed a new technique for detecting, recognizing, and modeling pole-like urban object from 3D point clouds. Given a 3D point cloud representing the cluttered urban scenes, the technique can rapidly localize the pole-like objects in the point clouds, segment the objects out from backgrounds, infer their geometric structures and shapes, classify them to semantic labels, and then construct their 3D models representing as point clouds or polygon surfaces.

Figure 4 illustrates the algorithmic pipeline of the developed system for detection and recognition of pole-like urban object from 3D point clouds. The input is a large-scale point cloud of an urban area. The outputs are detected the pole-like objects with their semantic label classifications in three categories: street lights, utility poles, and signs.

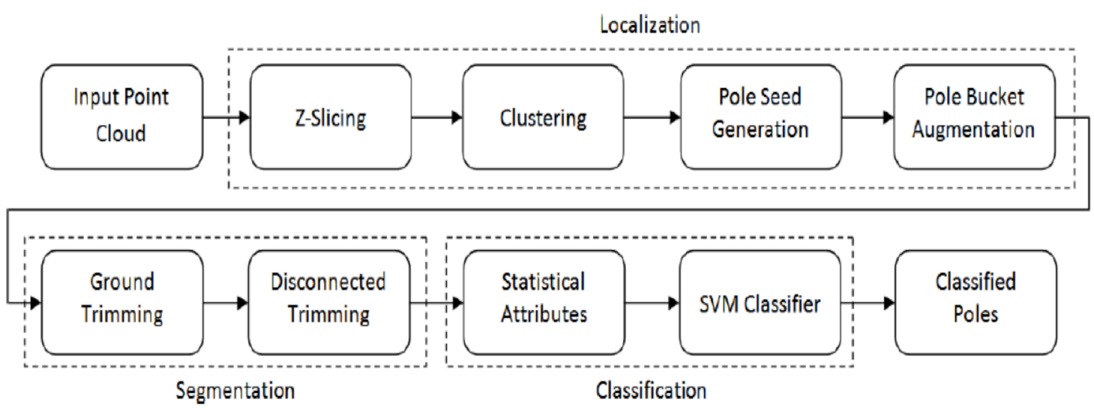

Figure 4: System architecture of pole-like urban object detection and modeling

Technically, there are three major stages of the processing. The first stage is localization, where all possible locations of pole-like objects are extracted from point clouds. In this stage, we make use of the unique geometry and functionality characteristics of pole-like objects to localize and extract the objects in the cluttered point clouds including the steps: point cloud slicing and clustering, pole seed generation, and pole bucket augmentation. The second stage is segmentation, in which the ground and other disconnected components are trimmed at the candidate locations. The method of 3D region growing is adapted to progressively segment out the objects from background. The segmentation results are a set of point clusters representing the locations and shapes of pole-like object candidates. Finally, several statistical and geometrical attributes are computed for each candidate using the extended distribution features to classify the candidates with a Support Vector Machine (SVM) classification mechanism.

We have evaluated the techniques with various datasets. Figure 5 shows the example results applied to the Ottawa datasets. More results of the work published in the premier forum - IEEE International Conference on Robotics and Automation (ICRA) in 2015.

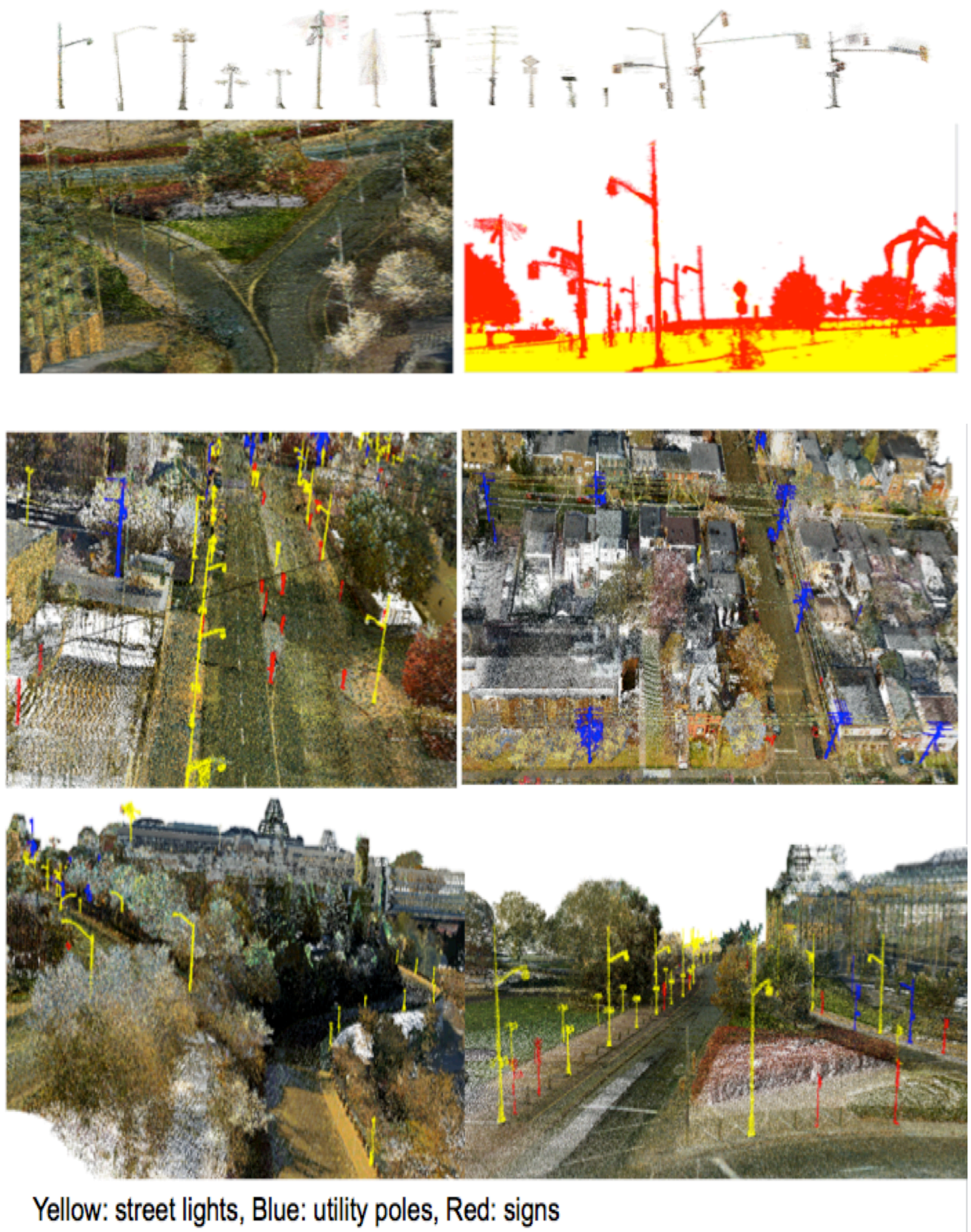

Figure 5: Detection and modeling results of Ottawa point cloud datasets

2.4. Deep Learning for Detection and modeling of vehicles

Vehicle is another class of commonly occurring object in urban sites. To obtain a complex scene description, we need a hierarchical modeling approach. At low-level individual vehicle has to be detected and modeled with accurate location, size and orientation estimation from various urban settings such as on the roads, in parking lots, or in front of the buildings, etc.

We developed a new technique for detecting, recognizing, and modeling vehicles from 3D point clouds. Uniquely, we extended the 2D deep Convolutional Neural Network (CNN) to the 3D point clouds. Deep learning is an emerging technique recently developed in machine learning community, and it has been successfully applied to resolve many 2D image classification problems under challenging conditions. Our work represents the state of the art of extending the deep learning technique to 3D point cloud processing.

Figure 6 illustrates the algorithmic architecture of the developed system for detection and recognition of vehicles from 3D point clouds. The input is large-scale urban point clouds. The outputs are detected individual vehicles with estimated locations, sizes, and orientations in the point clouds.

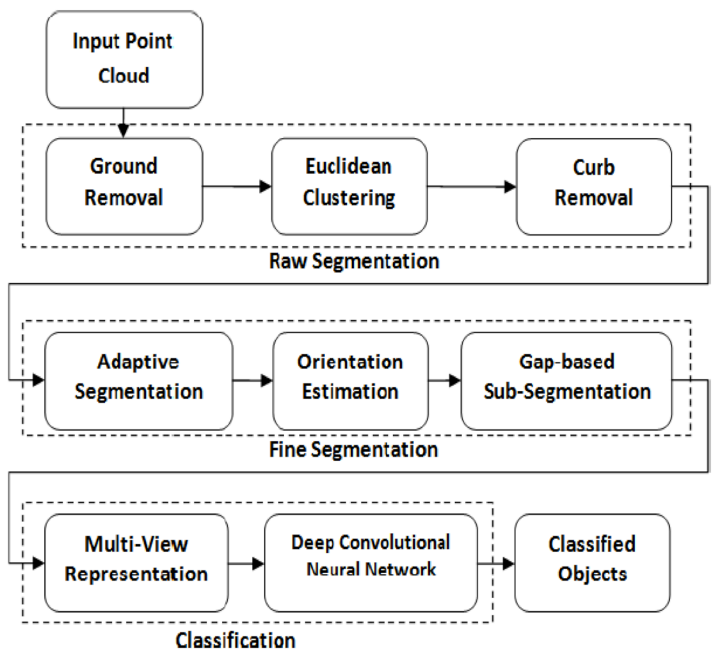

Figure 6: System architecture of CNN vehicles detection and modeling

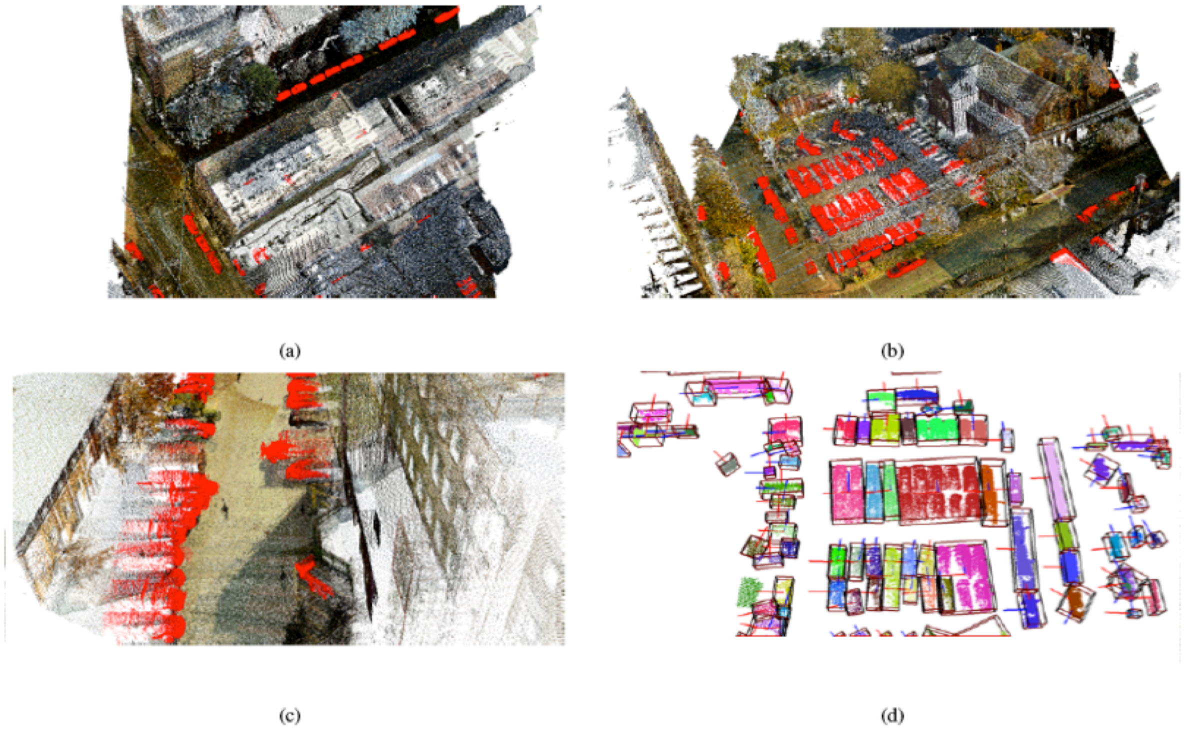

Figure 7: Detected vehicles in various scenes: (a) on the roads, (b) in parking lots, and (c) near the buildings. (d) shows the detected vehicles with estimated locations, sizes, and orientations

We evaluated the developed techniques with various datasets. Figure 7 shows the example results of detected vehicles in various scenes, i.e. on the roads, in parking lots, and near the buildings. This new developed work is prepared to submit for publication.

2.5. 3D CNN for Semantic Labeling of Point Clouds

We introduce a new framework of 3D Convolutional Neural Network (3D-CNN) and design effective algorithms for labeling complex 3D point could data. We present solutions for efficiently handling large data during the voxelization, training and testing of the 3D deep learning network.

The developed 3D CNN labeling system is depicted in Figure 8, which is composed of an offline training module and an online testing module.

The offline training takes the annotated training data as input. The training data are parsed through a voxelization process that generates occupancy voxel grids centered at a set of keypoints. The keypoints are generated randomly, and the numbers of keypoints are balanced across different categories. The labels of the voxel grids are decided by the dominating category in the cell around the keypoint. Then, the occupancy voxels and the labels are fed to a 3D Convolutional Neural Network, which is composed of two 3D convolutional layers, two 3D max-pooling layers, a fully connected layer and a logistic regression layer. The best parameters during training are saved.

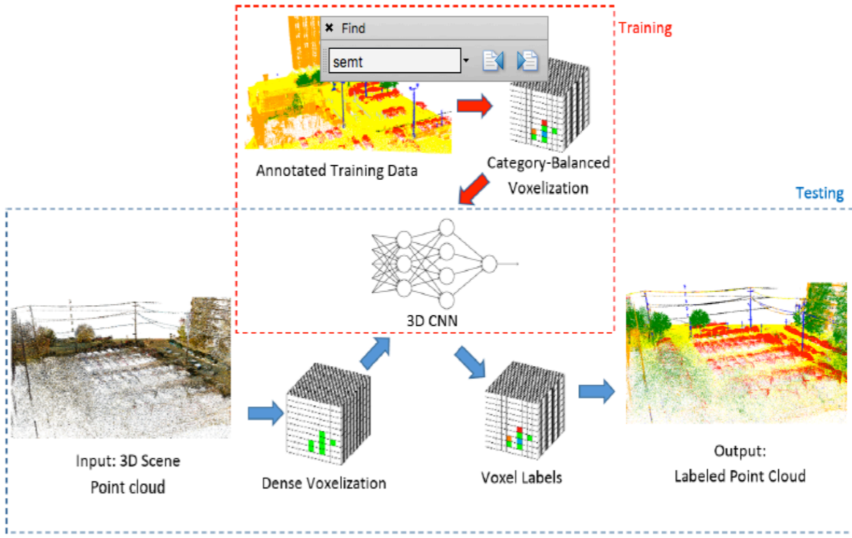

Figure 8: System architecture of 3D CNN Semantic Labeling

The online testing takes a raw point cloud without labels as input. The point cloud is parsed through a dense voxel grid and results in a set of occupancy voxels centered at every grid centers, respectively. The voxels are then used as the input to the trained 3D convolutional network, and every voxel grid would produce exactly one label. The inferred labels are then mapped back to the original point cloud to produce a pointwise labeling result.

We have evaluated the approach with various dataset. Figure 9 shows the confusion matrix across the categories. The overall precision for point labeling of all categories is 93.0%. Figure 10 shows the example results of labeling for the Ottawa dataset.

	Plane	Tree	Building	Car	Pole	Wire	Others
Plane	0.95	0.06	0.03	0.01	0.02	0.12	0.04
Tree	0.00	0.78	0.01	0.00	0.00	0.00	0.28
Building	0.02	0.06	0.89	0.00	0.03	0.02	0.02
Car	0.02	0.03	0.02	0.97	0.03	0.00	0.03
Pole	0.00	0.03	0.02	0.00	0.87	0.01	0.06
Wire	0.00	0.00	0.01	0.00	0.03	0.84	0.05
Others	0.00	0.03	0.01	0.02	0.02	0.00	0.51

Figure 9: Confusion matrix across the categories

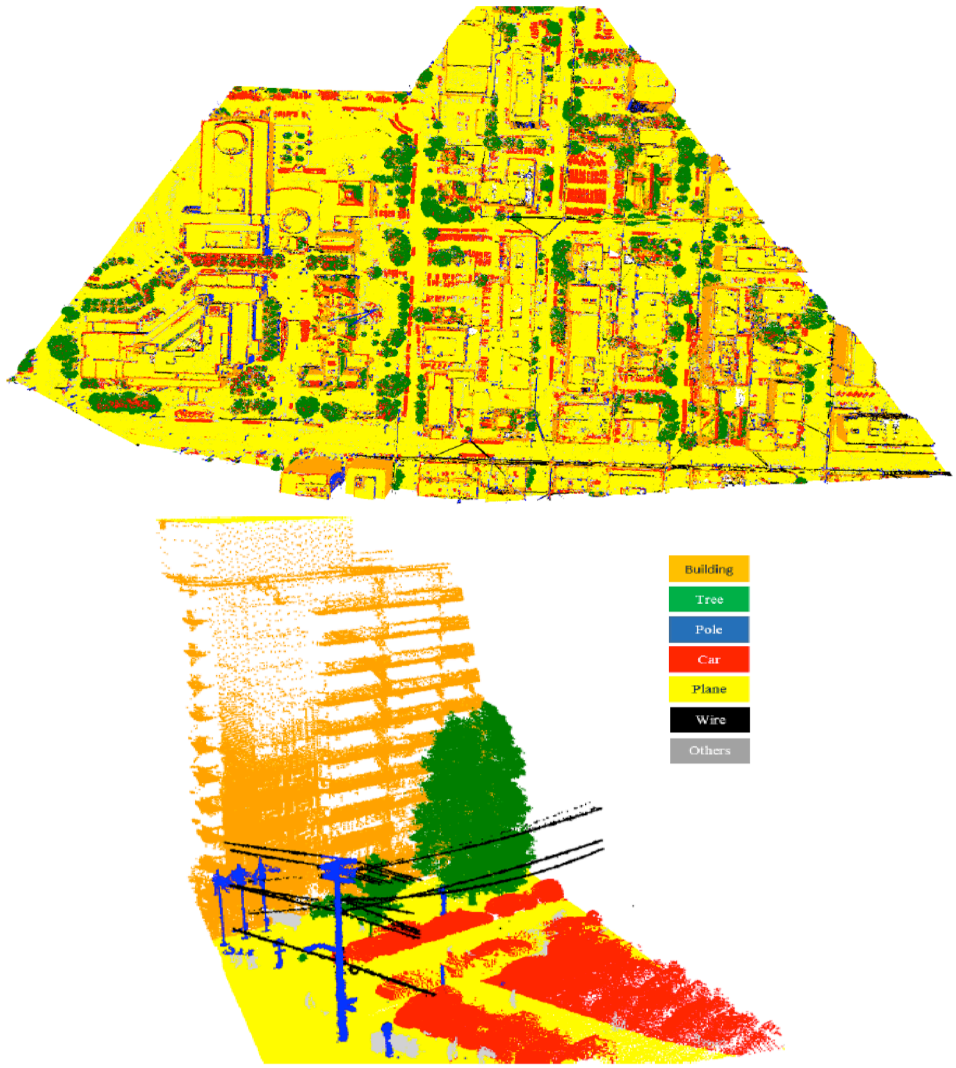

Figure 10: Example results of semantic labeling for the Ottawa dataset

2.6. System integrations

We integrated the developed components into MBR modeling system to reconstruct and model wide-area urban environments from point clouds. Given a 3D point cloud representing the cluttered urban scenes, the system automatically localizes the objects of interest in the point clouds, segment the objects out from backgrounds, classify them to the labeled targets in the database, and construct their 3D models representing as point clouds or polygon surfaces.

Currently, we have integrated the modules of building modeling, terrain modeling, pole-like object modeling, and vehicle modeling as illustrated in Figure 11, being able to reconstruct and model full-scale urban environments.

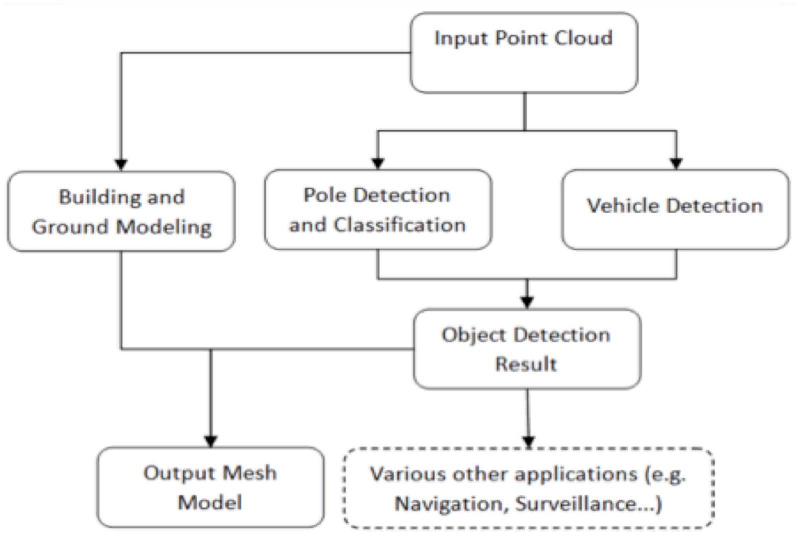

Figure 11: Integrated modeling system

Figure 12: Given 3D point clouds of urban scene (left), the system automatically reconstructs and creates full-scale 3D models of the scene including various large-scale objects and small-scale objects (right)

3. Publications and Patents

Publications:

- Jing Huang and Suya You, "Point Cloud Labeling using 3D Convolutional Neural Network," *International Conference on Pattern Recognition (ICPR)*, December 2016
- Jing Huang and Suya You, "Vehicle Detection in Urban Point Clouds with Orthogonal-View Convolutional Neural Network," *IEEE International Conference on Image Processing (ICIP)*, September 2016
- Jing Huang and Suya You, "Change Detection in Laser-Scanned Data of Industrial Sites", *IEEE Winter Conference on Applications of Computer Vision (WACV*), January 2015
- Guan Pang, Rongqi Qiu, Jing Huang, Suya You, and Ulrich Neumann "Automatic 3D Industrial Point Cloud Modeling and Recognition", *IAPR International Conference on Machine Vision Applications (ICMVA)*, May 2015
- Jing Huang and Suya You, "Detection and Recognition of Pole-Like Urban Objects from Point Clouds", *IEEE International Conference on Robotics and Automation (ICRA),* Seattle, WA, May 2015
- Jing Huang and Suya You, "Segmentation and Matching: Towards a Robust Object Detection System", *IEEE Winter on the Applications of Computer Vision (WACV)*, 2014, Steamboat Springs, CO, USA
- Wei Guan and Suya You, "Robust Image Matching with Line Context", *24th British Machine Vision Conference (BMVC 2013)*, 2013, Bristol, UK
- Jing Huang and Suya You, "Detecting Objects in Scene Point Cloud: A Combinational Approach", *International Conference on 3D Vision (3DV),* 2013, Seattle, USA

Patents and Invention Disclosures:

- Guan Pang, Jing Huang, Suya You, and Ulrich Neumann, "Multi-view 3D Object Recognition From a Point Cloud and Change Detection", US patent application, filed on March 2015
- Suya You, Jing Huang, Amir Anvar, and Christopher Fisher, "System and Method of Detecting Objects in Scene Point Cloud", US patent application, filed on 2013
- Ulrich Neumann, Suya You, "Three-Dimensional Point Processing and Model Generation", U.S. Patent Application, filled on 2013

4. Technology Transfer and Collaborations

We presented and demonstrated the developed techniques to several government agencies and defense contractors including: NGA, ISR, Lockheed Martin, and Northrup Grumman. Initial software transfer and evaluation for DoD and commercial applications is under discussion. We also presented the developed techniques to DoD's ISR program for wide-area surveillance applications with collected LiDAR data to be able to recognize and model targets of interest in noisy urban environments.

In cooperation with the Chevron Corporation, we have developed two application scenarios for petroleum industry: Rapid Modeling 3D Assets for Virtualization, and Smart Oilfield Safety Net.

5. Conclusions

In conclusion, we completed all the proposed tasks and achieved the anticipated results, which includes: Task-1: Model by recognition framework, Task-2: Robust 3D shape matching, and Task-3: Integrate/test/optimize. We developed and evaluated algorithms, showed results for simulation data and real world data, and provided measured performance. The techniques we developed and the results of that effort are very promising and make a compelling case that the developed MBR system and core techniques are in fact feasible and attractive as new approach to the challenging problems of wide-area urban modeling and object labeling.

We appreciate the generous supporting from ARO for this research and the project manager Dr. Joseph M Coyle.