RGC Ref.: N_HKUST607/11 NSFC Ref. : 61161160567
(please insert ref. above)

The Research Grants Council of Hong Kong
NSFC/RGC Joint Research Scheme
Joint Completion Report

(Please attach a copy of the completion report submitted to the NSFC by the Mainland researcher)

Part A: The Project and Investigator(s)

1. Project Title

Theory and Practice of Large-Scale 3D Urban Reconstruction and Modeling

2. Investigator(s) and Academic Department/Units Involved

	Hong Kong Team	Mainland Team
Name of Principal Investigator *(with title)*	Long QUAN	Baoquan CHEN
Post	Professor	Professor
Unit / Department / Institution	The Department of Computer Science and Engineering/ HKUST	Visual Computing Research Center/ Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences
Contact Information	The Department of Computer Science and Engineering/ HKUST	Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences
Co-investigator(s) *(with title and institution)*		

3. Project Duration

	Original	Revised	Date of RGC/ Institution Approval *(must be quoted)*
Project Start date	1 Jan 2012		
Project Completion date	31 Dec 2014		
Duration *(in month)*	36		

Deadline for Submission of Completion Report	31 Dec 2015		

Part B: The Completion Report

5. Project Objectives

5.1 Objectives as per original application

1. Development of a hybrid data fusion methodology from both images and laser scans systematically captured from a moving platform on the ground and in the air.

2. Development of a hybrid semantic segmentation approach, which is capable of automatically segmenting, recognizing and decomposing the large-scale input data at the object level.

3. Development of a hybrid image-based and point-based building modeling approach.

4. Development of a hybrid classification method which categorizes the input vegetation data into different species of trees in the given environment.

5. Development of a hybrid image-based and point-based vegetation modeling approach, and also development of a vegetation simplification approach to cope with the complexity of the vegetation models in the large-scale environment.

6. Development of an inverse procedural method that builds up large-scale urban environment from all available GIS data such as aerial photographs, street-view images, scans, and maps.

7. Pilot application of the approach to building some parts of the city of Hong Kong and Shenzhen.

5.2 Revised Objectives

Date of approval from the RGC: _____

Reasons for the change: _____

_____

6. Research Outcome

Major findings and research outcome
(maximum 1 page; please make reference to Part C where necessary)

In this project, we have developed a hybrid approach to fuse the visual image data and scanning data. The integrated data has also been semantically segmented based on some prior knowledges of the captured scenes. In our 2014 publication of Part C, we have proposed a method for the parsing of images and scanned point cloud in large-scale street environment. The proposed method significantly reduces the intensive labeling cost in previous works by automatically generating training data from the input data. The automatic generation of training data begins with the initialization of training data with weak priors in the street environment, followed by a filtering scheme to remove mislabeled training samples. We formulate the filtering as a binary labeling optimization problem over a conditional random filed that we call object graph, simultaneously integrating spatial smoothness preference and label consistency between 2D and 3D. Toward the final parsing, with the automatically generated training data, a CRF-based parsing method that integrates the coordination of image appearance and 3D geometry is adopted to perform the parsing of large-scale street scenes. The proposed approach is evaluated on city-scale Google Street View data, with an encouraging parsing performance demonstrated.

For the modeling properly speaking, point-based and image-based approaches are both developed to take advantage of the both available information. In particular, for vegetation data, a classification method has also been developed for better modeling of different types of observed vegetation data. Moreover, important simplification algorithms have been investigated and developed to make the 3D representations more efficient in navigation and applications. For facades, as there are available procedural rules which could be efficiently used to improve the modeling and reduce the complexity. In our 2013 publication of Part C, we have developed an image-based modeling of unwrappable facades. An unwrappable facade is represented by the mutually orthogonal baseline and profile. We first reconstruct semi-dense 3D points from images, then the baseline and profile are extracted from the

point cloud to construct the base shape and compose the textures of the building from the images. Through our unwrapping process, the reconstructed 3D points and composed textures are further mappedto an unwrapped space that is parameterized by the baseline and profile. In doing so, the unwrapped space becomes equivalent to the planar space in which planar fac¸ade modeling techniques can be used to reconstruct the details of the buildings. Finally, the augmented details can be wrapped back to the original 3D space to generate the final model. This newly introduced unwrappable representation extends the state of the art modeling for planar facades to a more general class of fac¸ades. We demonstrate the power of the unwrappable representation with a few examples in which the facade is not planar.

Potential for further development of the research and the proposed course of action *(maximum half a page)*

In the current project, with scanned point cloud of low resolution, like the Google Street View data we used in our experiment, it is difficult to extract small objects as the points on them are quite few, which hinders the automatic generation of training data for these categories. Therefore, currently, the targeted categories in the proposed method are constrained to the major categories in the street view. Given scanned point cloud captured by laser sensors with higher resolution, it is reasonable to expect that the proposed method can be generalized to more categories. Many research directions along this line of unwrappable representations remain. Many modern landmark buildings are of free-form surfaces which our representation is unable to describe. It is interesting to extend the idea unwrapping to handle such surfaces. For automatic modeling of unwrappable facades, it is the key to develop a technique to automatically decompose the point cloud into piecewise unwrappable components whenever needed. Joint 2D and 3D analysis in the unwrapped space is an interesting direction to further automate the modeling process. In order to increase the variation of the synthesis model, we plan to extract grammars about repetition from both images and geometries.

7. The Layman's Summary

(describe in layman's language the nature, significance and value of the research project, in no more than 200 words)

In this project, we have investigated an automated approach for the three-dimensional modeling of a large-scale urban environment from images. There are three major components. First, it primarily uses large amount of images captured by cameras or videos at ground level for building façades that are equipped with a detailed geometry and appearance suitable for viewing on the ground level as the real 'street view'; This complements and enriches the current 2D panorama-like street views deployed by Google. Second, it uses the aerial images and vector maps to generate the ground model and building roofs. Third, it uses both ground level images and aerial images to generate models of urban vegetation.

Part C: Research Output

8. Peer-reviewed journal publication(s) arising directly from this research project

(Please attach a copy of each publication and/or the letter of acceptance if not yet submitted in the previous progress report(s). All listed publications must acknowledge RGC's funding support by quoting the specific grant reference.)

The Latest Status of Publications				Author(s) ***(bold** the authors belonging to the project teams and denote the corresponding author with an asterisk*)*	Title and Journal/ Book *(with the volume, pages and other necessary publishing details specified)*	Submitted to RGC *(indicate the year ending of the relevant progress report)*	Attached to this report *(Yes or No)*	Acknowledged the support of this Joint Research Scheme *(Yes or No)*	Accessible from the institutional repository *(Yes or No)*
Year of publication	Year of Acceptance *(For paper accepted but not yet published)*	Under Review	Under Preparation *(optional)*						
2013				Tian Fang*, Zhexi Wang, Honghui Zhang, and Long Quan	Image-based Modeling of Unwrappable Facades IEEE Transactions on Visualization and Computer Graphics	30/06/2013	Yes	Yes	No
2014				Honghui Zhang*, Jinglu Wang, Tian Fang, and Long Quan	Joint Segmentation of Images and Scanned Point Cloud in Large-Scale Street Scenes with Low Annotation Cost IEEE Transactions On IMAGE PROCESSING (TIP 2014)	31/12/2014	Yes	Yes	No

9. Recognized international conference(s) in which paper(s) related to this research project was/were delivered *(Please attach a copy of each delivered paper. All listed papers must acknowledge RGC's funding support by quoting the specific grant reference.)*

Month/Year/ Place	Title	Conference Name	Submitted to RGC *(indicate the year ending of the relevant progress report)*	Attached to this report *(Yes or No)*	Acknowledged the support of this Joint Research Scheme *(Yes or No)*	Accessible from the institutional repository *(Yes or No)*
June 2012 USA	Parsing Façade with Rank-one Approximation	IEEE Computer Vision and Pattern Recognition 2012	30/06/2013	Yes	Yes	No
June 2014 USA	Local Readjustment for High-Resolution 3D Reconstruction	IEEE Computer Vision and Pattern Recognition 2014	31/12/2014	Yes	Yes	No
Nov. 2014 Singapore	Multi-view Geometry Compression	The 12th Asian Conference on Computer Vision (ACCV 2014)	31/12/2014	Yes	Yes	No
Nov. 2014 Singapore	Multi-scale Tetrahedral Fusion of a Similarity Reconstruction and Noisy Positional Measurements	The 12th Asian Conference on Computer Vision (ACCV 2014)	31/12/2014	Yes	Yes	No
Dec. 2013 Sydney	Learning CRFs for Image Parsing with Adaptive Subgradient Descent	IEEE International Conference on Computer Vision 2013	31/12/2014	Yes	Yes	No

10. Student(s) trained *(Please attach a copy of the title page of the thesis.)*

Name	Degree registered for	Date of registration	Date of thesis submission/ graduation
Honghui Zhang	PhD	Sep 2008	Aug 2012
Peng Zhao	PhD	Sep 2007	Aug 2012

11. Other impact *(e.g. award of patents or prizes, collaboration with other research institutions, technology transfer, etc.)*

N.A.