Free and Open Source Software for Geospatial (FOSS4G) Conference Proceedings

Volume 15 *Seoul, South Korea* Article 19

2015

Semi-Automated Building Footprint Extraction, Delineation And 3D Visualisation Of The University Of The Philippines Main Campus From LIDAR Data Using GIS-Based Open Source Software

Jara Kaye Villanueva
Laboratory Melchor Hall, College of Engineering University of the Philippines Diliman

Ma. Rosario Concepcion Ang
Laboratory Melchor Hall, College of Engineering University of the Philippines Diliman

Loureal Camille Inocencio
Laboratory Melchor Hall, College of Engineering University of the Philippines Diliman

Ma. Victoria Rejuso
Laboratory Melchor Hall, College of Engineering University of the Philippines Diliman

Recommended Citation

Villanueva, Jara Kaye; Ang, Ma. Rosario Concepcion; Inocencio, Loureal Camille; and Rejuso, Ma. Victoria (2015) "Semi-Automated Building Footprint Extraction, Delineation And 3D Visualisation Of The University Of The Philippines Main Campus From LIDAR Data Using GIS-Based Open Source Software," *Free and Open Source Software for Geospatial (FOSS4G) Conference Proceedings*: Vol. 15 , Article 19.
DOI: https://doi.org/10.7275/R5KD1W3F
Available at: https://scholarworks.umass.edu/foss4g/vol15/iss1/19

PO-09

SEMI-AUTOMATED BUILDING FOOTPRINT EXTRACTION, DELINEATION AND 3D VISUALIZATION OF THE UNIVERSITY OF THE PHILIPPINES MAIN CAMPUS FROM LIDAR DATA USING GIS-BASED OPEN SOURCE SOFTWARE

Jara Kaye Villanueva[1], Ma. Rosario Concepcion Ang[2],
Loureal Camille Inocencio[3], Ma. Victoria Rejuso[4]

Rm. 406 AGST Laboratory Melchor Hall, College of Engineering
University of the Philippines Diliman, Quezon City 1101 +6329818500 loc 3147
Email: [1]jarakayevillanueva@gmail.com, [2] moang@up.edu.ph,
[3]lcvinocencio@gmail.com, [4] victoria.rejuso@gmail.com

ABSTRACT

The advent of three-dimensional building footprint extraction and visualization has recently been explored due to its applications in urban planning, transportation, environmental monitoring, and modelling. Through the use of LiDAR (Light Detection and Ranging) data products such as Digital Terrain Model (DTM) and Digital Surface Model (DSM), useful information such as elevation, size, and shape can be obtained from the processed point clouds. Different studies regarding building extraction procedures have been conducted and explored. Algorithms have also been developed using LiDAR and GIS-based software that assume a rectangular form of polygons as building structures. As such, this study discusses an approach in developing a semi-automated building footprint extraction and three-dimensional visualization of the University of the Philippines (UP) main campus using solely LiDAR data and GIS-based open source software - QGIS Chugiak Version and GRASS 7.0. The UP-Diliman Campus situated in Quezon City, Philippines consists of various polygonal structured buildings which makes it a good subject area for this study. The proposed scheme composed of three major parts: LiDAR pre-processing, building footprint extraction and delineation, and three-dimensional building visualization. Normalized digital surface model (nDSM) can be generated using DTM and DSM. Under building footprint extraction and delineation, four parameters were set: height threshold, area scope, topographic modelling, and smoothing tolerance. For the classification by height, a threshold is set to remove structures with height less than the predetermined value of elevation. Classification by area removes the objects with areas that do not fall within the predetermined area scope. Topographic modelling is used mainly to separate building from other entities. Smoothing tolerance simplifies building outlines. The accuracy assessment for the extraction and delineation is quantified using ground truth data in comparison with the extracted polygons. As compared to manual digitization of building polygons, this semi-automation can be more efficient in extracting and delineating building footprint in areas with large scope. Experimental results indicate that the proposed scheme provides a promising solution for 3D building extraction and delineation using LiDAR data processed in QGIS Chugiak Version and GRASS 7.0. This methodology is a valuable tool for urban planning and modelling.

1. INTRODUCTION

1.1 Background of the Study

Three-dimensional modelling can be useful in urban management, planning, and monitoring as compared to two-dimensional platform.For building delineation and extraction and the generation in three dimensional space, its automation reduces the time and cost it entails as compared to manual process. Several studies have been conducted with its focus on both the automation and semi-automation of this process using high-resolution image data.

In recent years, Light Detection and Ranging or LiDAR Technology has recently been explored due to its promising applications in environmental planning (Perén*et al* 2015), urban management(Coirier*et al* 2006), and energy prediction (Kalmikov*et al* 2010). This high-resolution data contain useful information for building extraction such as height, shape, and size. Using spatial tools in Geographic Information System or GIS, the analysis and visualization of data can be accomplished. Free and open source GIS software such as QGIS and GRASS has been widely utilized due to its increasing security against malware and decreasing cost for software acquisition.

At an appropriate level of detail, this paper discusses a semi-automated methodology for building footprint extraction, delineation, and 3D visualization of the University of the Philippines Main Campus from LiDAR data using QGIS 2.4 and GRASS 7.0 as the GIS-based open source software.For urban settings with a considerable large area, the semi-automation of extraction and visualization methods can greatly reduce the processing time and cost.

1.2 Objectives of the Study:

This research aims to develop a semi- building footprint extraction, delineation and three-dimensional visualization of the University of the Philippines Main Campus from LiDAR data using GIS-based open source software through the utilization of LiDAR data products such as Digital Terrain Model (DTM) and Digital Surface Model (DSM).

1.3 Scope of Study:

This study covers a portion of University of the Philippines Diliman Campus in Quezon City with an area of .7782 square kilometer. The area includes buildings with complex geometries and tree canopies which make it an ideal test site for building extraction methodology.The GIS software used is QGIS version 2.4 with the help of GRASS 7.

Figure 1. The image shows the portion of UP Main Campus used in this study

For the 3D visualization, a QGIS plug-in created by Minoru Akagi was used. Three2js is a QGIS plug-in that uses 3D library repository and WebGL-enabled browser for visualization. The parameters and their corresponding threshold used is dependent on the study area as a result of trial and error and the researchers' judgement and skill.

1.4 Overview of Related Works

Several algorithms for building extraction and 3D visualization using different methods have already been developed. This can be manual building detection, semi-automated/automated building detection using Orthoimages or LiDAR data, and semi-automated/automated delineation using both data. A brief overview of related works is discussed in this chapter.

Previous studies developed the use of solely orthoimages for urban feature extraction (Ming *et al* 2005, Wei *et al* 2005, and Varshney*et al* 2011). Detection of building patterns are done through parcel units. Parameters did not include height information which explains why LiDAR image was not used. High resolution LiDAR images are usually accompanied by RGB data in the extraction of urban features. Studies by Jiang (*et al* 2008) and Syed (*et al* 2005) used RGB data to get the shape, spectral, and semantic information of the objects through multi-resolution segmentation. The height information contained in DSM was used to distinguish buildings from ground. Both adopted the concept of object-oriented classification.

However, high resolution images like LiDAR and RGB data are not readily accessible and propriety GIS software for spatial processing costs a lot. For this study, the use of solely LiDAR data with the use of spatial tools embedded in Free and Open Source Software (FOSS) to delineate, extract, and visualize urban geometries in three-dimensional space is highlighted.

2. METHODOLOGY

2.1 Process Workflow

The workflow as presented in Figure 2 is the proposed scheme and is composed of three main processes. The first step is the pre-processing of LiDAR data products DTM and DSM. Using nDSM as the result from the first step, building delineation and extraction is done by using height, slope, and area thresholds. This building footprint is then modelled into three-dimensional space for visualization. Further discussions on each process can be found in the succeeding parts of methodology.

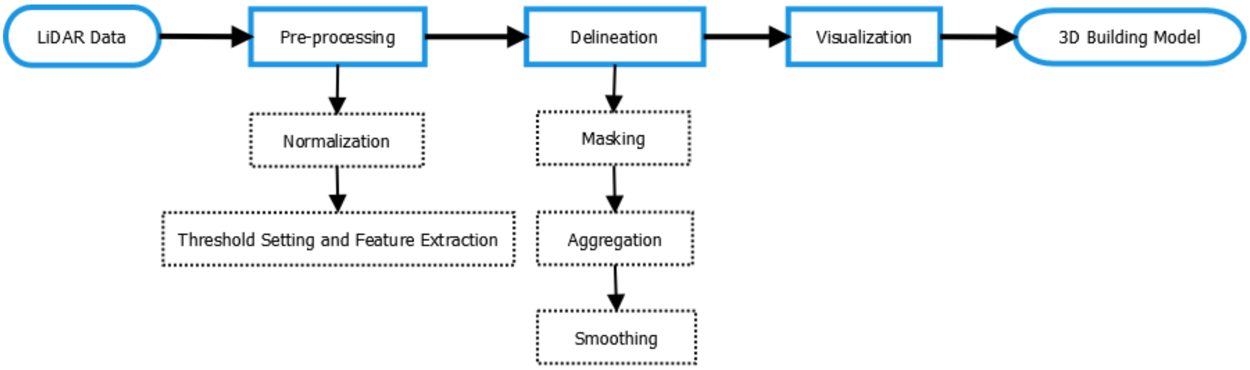

Figure 2. Workflow of the semi-automated building extraction and visualization

2.2 Test Area, and Datasets and GIS Tools Used

The data used are derived LiDAR products which include DTM and DSM. Both are acquired from National Mapping and Resource Information Authority's (NAMRIA) project, Collective Strengthening of Community Awareness on Natural Disasters (CSCAND) agencies. Both LiDAR data products have an image resolution of 1m x 1m with an uncompressed size of 3.03 MB. The area covers a portion of University of the Philippines Diliman Campus in Quezon City, measuring .7762 square kilometer.

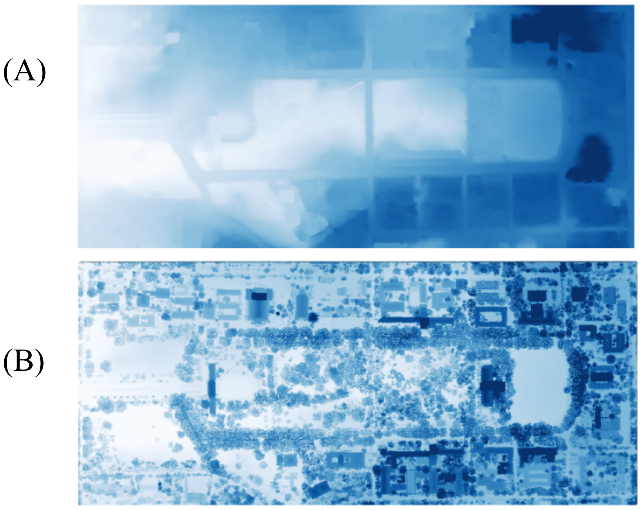

Figure 3.LiDAR data products used include (A) Digital Terrain Model (DTM) and (B) Digital Surface Model (DSM)

The semi-automation can be achieved through Geographic Information System (GIS) tools. GIS is a set of tools that analyze, store, interpret, process, and visualize all types of geospatial information. A couple of free and open source GIS software have recently been developed. Two of these free and open source GIS tool used in the study are QGIS 2.4 and GRASS 7.0. Quantum GIS or QGIS and GRASS are both free and open source software that can create, edit, analyze, and visualize geographic data. They are operational in various operating systems such as Windows, Mac, and Linux.

2.3. Processing

2.3.1. Pre-processing

The initial process consists of two processes: normalizing the Digital Elevation Model and threshold setting through trial and error.

2.3.1.1. Normalization

In LiDAR pre-processing, the concept of normalized digital surface model or nDSM is used. It can be obtained from Digital Terrain Model from Digital Surface Model. DTM delineates the elevation information of ground data or the terrain, while DSM contains the elevation information of surface features. Thus,by subtracting DTM from DTM, we will obtain the nDSM which represents the z data of surface features.

2.3.1.2. Threshold Setting and Feature Extraction

Feature extraction involves the defining of certain thresholds and parameters to differentiate and separate one feature from another. For delineation and extraction of building footprint, certain parameters are set. These include height, slope, and area. For each parameter, the optimal threshold is obtained. The values for the threshold are entirely dependent on the area. Thus, there is no constant threshold values applicable to all test sites. Finding the parameter values can be done through trial and error. The following parameters are used to delineate building from other features:

2.3.1.2.a. Height

To separate high features from ground, the initial parameter considered is height. In an urban setting, buildings have generally higher Z information than other classes. By applying the height threshold, initial building mask can be made. The high features extracted may still classes other than buildings such as vegetation and other objects. Through trial and error, the height threshold set is 2.5 meters.

2.3.1.2.b. Slope

Slope is the change in elevation and can be expressed in percent or in grade. The slope of the building is generally lower since it is planar, as compared to vegetation features such as trees.For this area, the threshold is set to 50 percent.

2.3.1.3.c. Terrain Ruggedness Index (TRI)

Terrain ruggedness index (TRI) describes the shape or roughness of surface which are derived from land cover. TRI can be determined through different methods. It can also be done through computing the standard deviation of slope or elevation. For this particular algorithm, the mean difference between eightneighbouring cells and center pixel is computed. Urban geometries have lower ruggedness index as compared to vegetation. Thus, this can be a variable threshold between buildings and trees.

2.3.1.2.d. Area

The area parameter setting is done after aggregation of delineated building polygons. Polygons with least areas are separated. This is to delete the small ungrouped polygons. The polygon area values used to segregate buildings from other polygons may vary from the largest to smallest area depending on the sizes of features. Visual judgment is used in this particular procedure.

2.3.2.Delineation

The entire extraction process consists of masking, aggregation, and smoothing of polygons. Through spatial tools in QGIS 2.4 and GRASS 7.0,of values not within the threshold set. GRASS 7.0 is used in getting the normalized DSM and the slope. In order to extract the designated attributes as previously defined in threshold setting, functions in QGIS such as Polygonize tool and Extract by Attribute tool is used. The Terrain Ruggedness Index (TRI) can be computed from GDAL plug-in.

2.3.3. Visualization

QGIS has a website for a repository of different plug-ins. For visualization of delineated and extracted buildings in three-dimensional space, Qgis2threejs plug-in was used. Qgis2threejs was created by Minoru Akagi. It is used to export map layers and building geometries into a WebGL-capable browser. This plug-in uses Three.js library is a lightweight 3D library that handles 2D graphics and renders it into three-dimensional space.

2.3.3.1. Data Overlaying

The user can use more than one layer of image data. The plug in reads and allows both raster and vector data types. Two data overlays were used in urban visualization. A screenshot of UP Diliman map from Google Maps is georeferenced in accordance with the projection of the building shapefiles. The second layer is the vector file of the building footprint containing height information.

2.3.3.2. 3D Modeling

Using the height information in the delineated building polygons, extrusion is done using the Three2jsplugin. In addition to that, the building polygons were also classified according to their Z information for visualization purposes in two-dimensional space.

3. RESULT AND DISCUSSION

The results from the semi-automated building footprint extraction, delineation, and 3D visualization of UP Diliman using QGIS and GRASS spatial toolswere presented below.The result of the difference between Digital Terrain Model and Digital Surface Model is shown below. Normalized Digital Surface Model (nDSM) contains the height information of features. Darker areas correspond to higher elevation values.

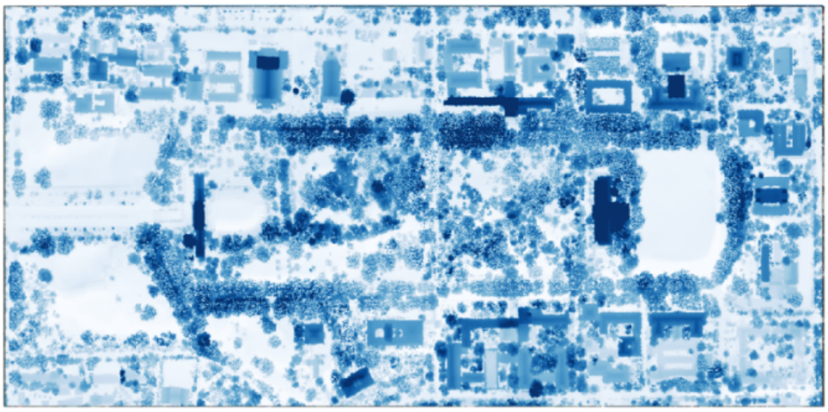

Figure 4. The image shows the normalized DSM or nDSM

Height threshold parameter is set to 2.5 meters such that high features and ground will be separated. These high features, however, still include buildings and trees.

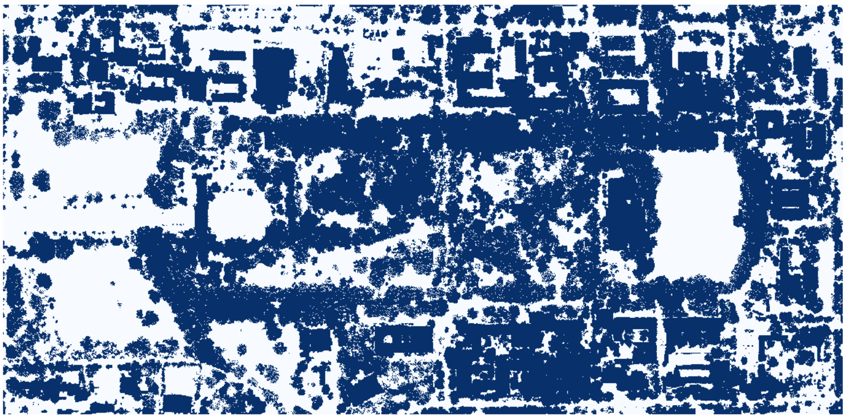

Figure 5. Image shows the result of the application of height threshold

To delineate urban geometries from vegetation, the slope and terrain ruggedness index parameters are used. Darker areas have higher values of slope and tri. As seen from the figures, the trees and edge of buildings have darker tones which indicate these have greater slope and TRI.

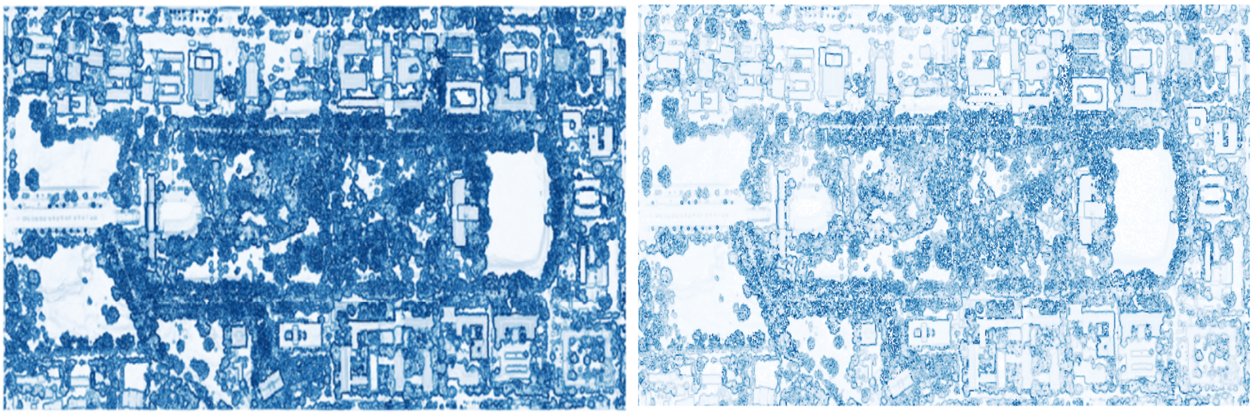

Figure 6. Left image shows the slope image and right corresponds to the TRI of the area

Using these thresholds, masking features other than urban geometries are done. Buildings are more delineated in this process. Some small polygons were eliminated through creating a threshold for area.

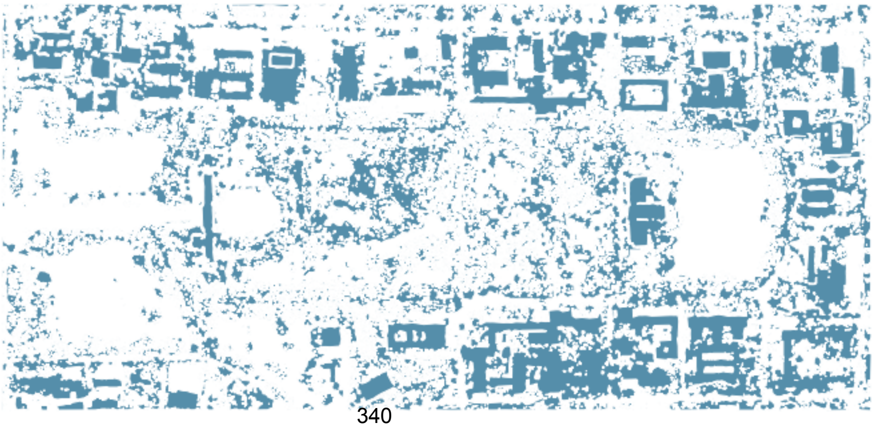

Figure 7. Result of the application of threshold parameters and masking

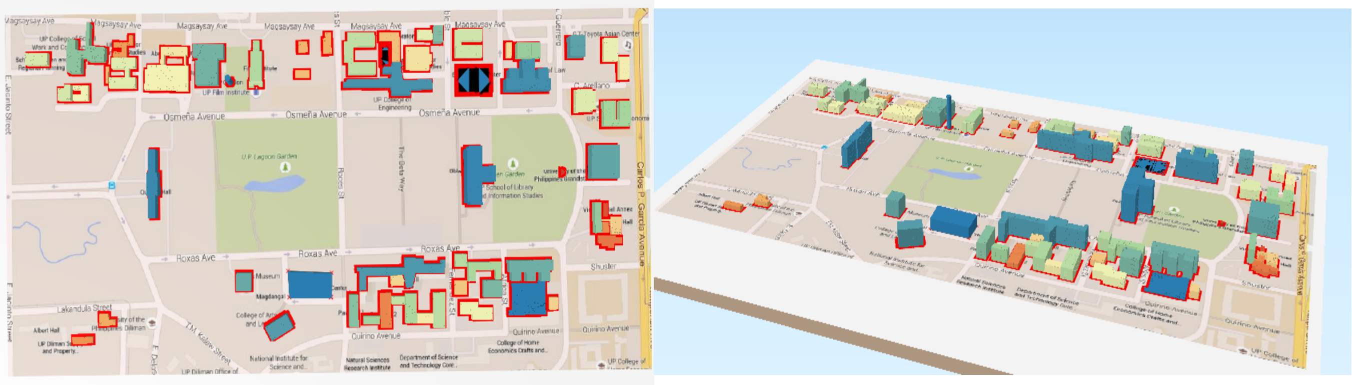

Figure 8. Visualization of building polygons using Qgis2threejs plug-in

3.2 Validation and Accuracy Assessment

In quantifying the quality of the result, comparison between the output of the semi-automated and manual detection was done.The process for validation is adopted from a quantitative evaluation of building detection as devised by Chungan Lin and RamakanNevatia.

The data from Open Street Map (OSM) were used for validation. These OSM shape files can be downloaded from their website and are readily available for public's use.

Figure 9. Open Street Map shapefiles in comparison to result of the study

One of the parameters used is detection percentage. An urban feature is considered detected even if only a fractional part of it is delineated. The corresponding formula is shown below:

$$\text{Detection Percentage} = \frac{(100\ X\ TP)}{(TP+TN)} \qquad (1)$$

Detection percentage quantifies the proportion of delineated building in comparison to the actual data or the data that can be obtained through manual digitization. Where:

TP = True Positive, urban feature detected by both manual and semi-automatic method

TN = True negative, urban feature detected by manual and not by semi-automatic

Table 1. Quantitative evaluation of the extraction result

True Positive	55 polygon buildings
True Negative	19 polygon buildings
Detection Percentage	74.324%

4. CONCLUSION AND RECOMMENDATION

This study utilized the use of Free and Open Source Software in the semi-automation of building footprint delineation, extraction, and visualization in UP Diliman. Particularly, the scheme made use of GIS spatial tools embedded in QGIS and GRASS for the analysis and modelling.

All the parameters and threshold used are based from the researchers' judgment and skill. Since these values are not constant, they are entirely dependent on the area of study. These can be determined through trial and error. These variables may differ depending on the type of urban setting. For locations with fairly less trees, cleaning up of unclassified polygons would be less, considering the height threshold that was initially set. However, for urban areas with buildings almost covered with trees, extraction of building footprint is more difficult.

This semi-automated process of analysis and visualization, however, is limited to urban features with flat roof. The current algorithm does not take into account the aspect and slope of the roof. Thus, for future studies and application which requires high accuracy, computation and modeling of shape and slope of roof is recommended. Semi-automation in batch process can be further developed in FOSS through python scripting or creation of a QGIS plug-in. Model builders can also be useful for simultaneous utilization of spatial tools.

In conclusion, for considerably large areas, this method can be less time consuming and more cost efficient as compared to manual digitization of building shapefiles and using proprietary software. Free and open source GIS software like QGIS and GRASS have a significantly great potential for the analysis, manipulation, and modelling of microscale regions particularly the urban areas.

5. References

Teo, T., Rau, J., Chen, L., Liu, J., & Hsu, W. (n.d.). Reconstruction of Complex Buildings using LIDAR and 2D Maps. *Lecture Notes in Geoinformation and Cartography Innovations in 3D Geo Information Systems,* 345-354.

Jiang, N., & Lin, X. G. (n.d.). Object-Oriented Buinding Extraction ByDsm and Very High-Resolution Orthoimages. *Archives*.

Ming, D. M. D., Luo, J. L. J., Shen, Z. S. Z., & Li, J. L. J. (2005). Features based parcel unit extraction from high resolution image. *Proceedings. 2005 IEEE International Geoscience and Remote Sensing Symposium, 2005. IGARSS '05.*,*3*(section 3), 5–8. http://doi.org/10.1109/IGARSS.2005.1526374

Syed, S., Dare, P., & Jones, S. (2005). Semi-automatic 3D building model generation from lidar and high resolution imagery. *Proceedings of SSC Spatial Intelligence*, 1421–1428. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:SEMI-AUTOMATIC+3D+BUILDING+MODEL+GENERATION+FROM+LIDAR+AND+HIGH+RESOLUTION+IMAGERY#0

Wei, W. W. W., Chen, X. C. X., & Ma, A. M. A. (2005). Object-oriented information extraction and application in high-resolution remote sensing image. *Proceedings. 2005 IEEE International Geoscience and Remote Sensing Symposium, 2005. IGARSS '05.*,*6*(C), 5–8. http://doi.org/10.1109/IGARSS.2005.1525737

Varshney, A., Rajesh, E., Sarabandi, P., & Solutions, R. M. (1962). Advanced region – based pattern recognition algorithms for building footprint extraction- an automatic approach to generate building inventory.

GIS GPS Geodesy Bookmark: Free Open Street Map data (i.e.shapefiles etc.) for the Philippines. (n.d.). Retrieved December 15, 2014, from http://gisgpsgeodesy.blogspot.com/2011/04/free-road-data-shapefiles-for.html

GIS 4 Geomorphology. (n.d.). Retrieved August 15, 2015, from http://gis4geomorphology.com/roughness-topographic-position/

Shridhar J., Panditrao S., and Luis A., "Airborne LiDAR and High Resolution Satellite Data for Rapid 3D Feature Extraction." National Remote Sensing Centre Indian Space Research Organisation. Web. 4 May 2015.

Shen, W. (2008) Building Boundary Extraction Based on LiDAR Point Clouds Data. The /International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B3b. Beijing

"Automated Building Extraction and Reconstruction from LIDAR Data." (2000) Geographic Resources Center. Web. 14 May 2015.

Dy, E. (2013) Automatic Generation of 3D Buildings and Trees from DSM and DEM Data using GIS. Print.

Vosselman G., and Dijkman, S., (2001) 3D Building Model Reconstruction from Point Clouds and Ground Plans. IAPRS, Vol. XXXIV, Part 3/W4, pp.37-43

Demir, N., and Baltsavias E., (2012) Automated Modeling of 3D Building Roofs Using Image and LiDAR Data. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences.

Shiravi S., Zhong M., Beykaei S., (2012) Accuracy Assessment of Building Extraction using LiDAR Data for Urban Planning/Transportartion Applications. 2012 Conference of the Transportation Association of Canada Fredericton, New Brunswick.

Coirier, W. J., & Kim, S. (2006, February). CFD modeling for urban area contaminant transport and dispersion: Model description and data requirements. In *Sixth Symposium of the Urban Environment, Atlanta, GA*.

Kalmikov, A., Dupont, G., Dykes, K., & Chan, C. (2010, May). Wind power resource assessment in complex urban environments: MIT campus case-study using CFD Analysis. In *AWEA 2010 WINDPOWER Conference. Dallas, USA*.

Perén, J. I., van Hooff, T., Leite, B. C. C., &Blocken, B. (2015). CFD analysis of cross-ventilation of a generic isolated building with asymmetric opening positions: impact of roof angle and opening location. *Building and Environment*,*85*, 263-276.