BUILDING MODELING FROM LIDAR AND AERIAL IMAGERY

Jinhui Hu, Ph.D. Student
Suya You, Research Assistant Professor
Ulrich Neumann, Associate Professor, Director
Kyung Kook Park, Ph.D. Student

CGIT, IMSC,
University of Southern California
3737 Wattway, PHE404
Los Angeles, California, 90089
{jinhuihu, suyay, uneumann, kyungkpa}@ graphics.usc.edu

Abstract

This paper presents a modeling system using airborne LiDAR and aerial imagery. Our approach is a hierarchical technique that allows users to create a hierarchical building model composed of geometric primitives. Linear primitives and high-order surface primitives are used for model fitting and refinement. To improve accuracy and efficiency, we use image information to aid model and refine processes. Both the knowledge-level and pixel-level information are used. The texture and color information from aerial image is used to automate the segmentation process. Building shape cues from range image are used to reduce the number of model hypotheses and computation complexity. Edges from high resolution aerial images are used to improve the model accuracy. We demonstrated the system's flexibility and capability for modeling wide range of complex buildings.

1 Introduction

3D urban models have many applications in urban planning, environment monitoring, geo-information systems, traffic managements, utility services, and military operations. In most of these cases the models of buildings, terrain features, and vegetations are the primary features of interest. Although urban models are useful for the reasons stated above, the creation of detailed wide-area models remains at best a difficult and time-consuming task [Hu, 2003; Ribarsky, 2002]. A wealth of research, employing a variety of sensing and modeling technologies, has been conducted to create detailed building models from imagery or from laser sensing data. Photogrammetry offers a cost-effective means to obtaining large-scale urban models. The techniques in this category use 2D images without any a priori 3D data. Different image sensors lend themselves to modeling systems developed for terrestrial or aerial images [Lee, 2000]. Recently, airborne LiDAR (Light Detection and Ranging) has become a rather important information source for generating high quality 3D digital surface models. A LiDAR sensor system permits an aircraft flyover to quickly collect a height field for a large environment with an accuracy of centimeters in height and sub-meter in ground position. Multiple passes of the aircraft are merged to ensure good coverage. Due to its advantages as an active technique for reliable 3D determination, LiDAR offers a fast and effective way to acquire models for a large urban site [Ahmed, 2002; Zhao, 2000].

While different sensors provide varied data for scene modeling, each of these data sources and corresponding techniques has their own advantages and disadvantages. Images proved detailed texture and color information and they can provide very high accuracy, making them necessary for texture data and appealing for extracting detailed model features. On the other hand, LIDAR data samples are dense 3D samples of building and terrain surfaces. A natural conclusion is to fuse these data sources to obtain more accurate and automatic urban models. One single sensor technology seems unlikely to produce detailed and varying characteristics of building models. Combining the geometry, photometry, and other sensing sources can compensate for the shortcomings of each sensing technology, and appears to be a promising methodology. This is main motivation for our work. In this paper, we address the issue of combining LiDAR and aerial imagery for rapid creation of accurate building models. The basic tenets of this work are: features extracted from high-resolution image can improve the accuracy of low-resolution LiDAR model features; and cues from LiDAR can aid in image segmentation, significantly reducing the computational complexity and processing time as well as improving the quality of model results.

This paper presents our recent extensions to a complete modeling system [You, 2003]. LiDAR is used to acquire build models for large urban areas. However, sample-rate limitations and measurement noise obscures small details and occlusions from vegetation and overhangs lead to the LiDAR model voids in many areas. Points on building edges and surfaces have to be segmented accurately from the ragged LiDAR model. Our approach employs several morphological filters operating on the LiDAR range data, and texture and color from aerial imagery to segment the targeted objects from background. To model the extracted 3D mesh model to produce constrained CG models, we present a primitive-based model refinement approach. Based on the shape of building rooftop, we classify a building section into one of several groups, and for each group we define a set of appropriate geometry primitives, including standard CG primitives and high-order surface primitives, fitting to the building's mesh data to represent the complete building structure.

The rest of the paper is organized as following: section 2 presents our primitive-based modeling system, section 3 discusses the approach of integrating aerial and range images to the modeling processes, section 4 presents our experimental results, and concludes the paper.

2 A primitive-based modeling system

2.1 System Overview

Our modeling system (Figure 1) begins with a model reconstruction phase followed by a model refinement and optimization phase. The model reconstruction phase processes raw LiDAR point cloud to create a regular-grid 3D mesh model of scene. Geo-referencing, data re-sampling, hole-filling, and tessellation comprise this phase.

The model refinement and optimization phase consisting of building extraction, model fitting, and refinement components processes the reconstructed 3D mesh model to create hierarchical building models. The global building footprints provided by the LiDAR and aerial imagery are used to determine the locations of buildings and extract them from surrounding terrain. Based on the shape of a building rooftop, we classify a building section into one of several groups, and for each group we define a set of appropriate geometry primitives, including standard CG primitives and high-order surface primitives. Once a building is extracted, the geometry primitives are iteratively fit to the building's mesh model data, and the best fitting models represent the complete building structure. The model refinement is a hierarchical approach that allows users to create a hierarchical building model composed of geometric primitives. This approach has demonstrated its flexibility and capability for wide range of complex buildings with irregular shapes.

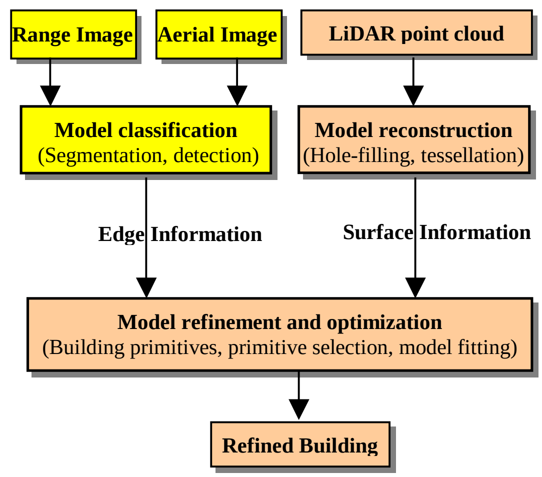

Figure 1. Algorithmic structure and work flow of proposed modeling system.

2.2 Primitive-based model refinement

According to the shapes of building rooftops (flat-roof, slope-roof, dome-roof, gable-roof, etc.), a set of geometric primitives are defined. They are linear fitting primitives: plane, cube, wedge, cylinder, polyhedron, and sphere, and nonlinear fitting primitive: superquadrics. These geometric primitives are then fit to the local mesh data to represent the local structures of building sections. Finally, the fitted local models are assembled to a complete building model.

Linear Primitive Fitting. Linear primitive models include: plane, cube, wedge, cylinder, polyhedron, and sphere. The parameters of these primitive models are estimated using linear least square fitting techniques. To fit a primitive model into a reconstructed mesh model, there are two types of parameters need to be estimated, i.e. edge parameter and surface parameter. So, the fitting process also includes two steps: edge fitting and surface fitting. In the following we detail a cube primitive fitting process as an example. A complete algorithm description for every type of primitives can refer to reference [You, 2003].

Based on two user initialized diagonal points, the approach automatically estimates all four corners of the cube roof by using a global direction parameter. Then these initial corner estimates are used to search for edge points based on shape connectivity. Many incorrect edge or noise points will be included if we only use the connectivity rule. To further refine the edge points, a depth filter and a slope filter are employed. Results showed that these two filters are very efficient to find correct edge points. Least square fitting technique is then used to estimate the best parameters for the edges of buildings. Similar to this edge points segment, surface points are also extracted using the depth and slope filters, as well as the parameters being estimated using least square fitting approach.

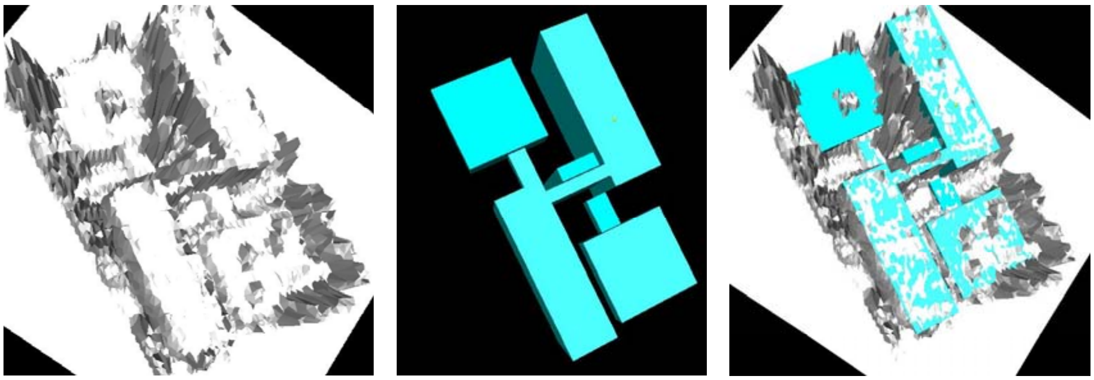

Figure 2. A complex building is modeled using multiple cuboid primitives. From left to right, original LiDAR model, refined model, and refined model embedded into original mesh.

Non-Linear Primitive Fitting. Quadric and high order curved surfaces cannot be modeled using the above linear primitives. To handle complex buildings with curved surfaces, non-linear primitives are introduced. Superquadrics are extensions of non-linear generic quadric surfaces. They have capability of describing a variety of curved shapes with a small number of parameters. We use the superquadric as a general form to model all the nonlinear high-order surfaces. The Levenberg-Marquardt (LM) method is used to perform the fitting process. Three steps comprising the model fitting approach are described as following.

Object segmentation: the region-growing approach is employed to segment the irregular object from its background. Given a seed-point, the algorithm automatically segments the seeded region based on a defined growing rule. In our implementation, the surface normal and depth information are used to supervise the growing procedure.

Initial surface fitting: to guarantee a converged optimal solution, an appropriate initial value is required for the LM algorithm. A sphere primitive fitting is used for system initialization.

High-order surface filling: once initialized, the system fits the ellipsoid primitive to the segmented surface points using LM algorithm. The algorithm typically needs several hundreds iterations (606 iterations in the example shown in Figure 3) to converge to a correct solution. Figure 3 shows the result of applying this approach to model the Los Angeles Arena.

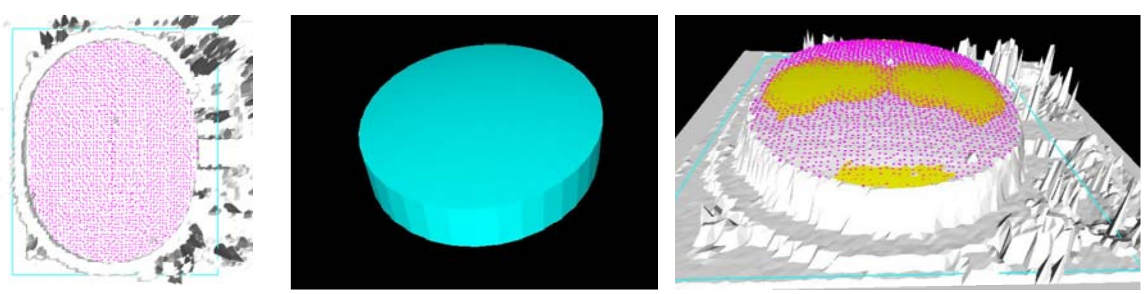

Figure 3. High-order primitive is used for modeling Los Angeles Arena: (a) Segmented edge and surface points (purple), (b) refined model with the ellipsoid primitive, and (c) the refined model embedded in the original mesh

3 Integration of aerial and range image for building modeling

3.1 Image registration

As mentioned before, 3D modeling can benefit from the integrated using of multiple sensor sources. These different data sources must first be spatially co-registered. The information extracted from one data source which is used to help the analysis of another data source can be roughly classified into two types: knowledge-level information and pixel-level information. The knowledge-level information is higher-level cue, e.g. a region information of building, or color and texture information of imagery. [Huertas, 2000] extracted shape cues from ISFAR images, and then used these cues to guide analysis of the EO panchromatic (PAN) images. Their results showed that combining different data sources in knowledge-level can significantly reduce the computation complexity of fusion algorithm. The pixel-level information is lower-level, such as edges, footprint, and corners. To use the pixel-level information, the registration error should be less than the minimum resolution of data sources. For example, if the resolution of a LiDAR data is 1 meter, and that of a aerial image is 0.5 meter, then the registration error should be less than two pixels. Otherwise it makes no sense to use information such as edges from images to improve the accuracy of LiDAR models.

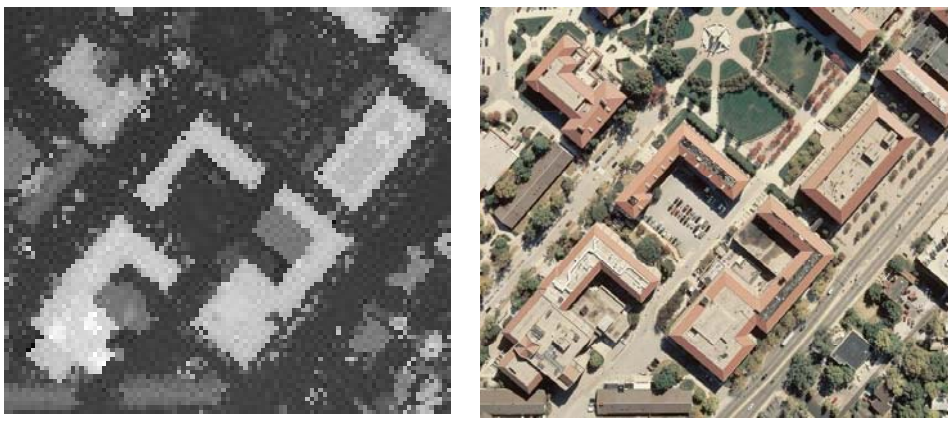

Figure 4. a) Low quality range image from LIDAR data. b) High resolution aerial image

Figure 4 shows a LiDAR range image of Purdue campus, and associated high-resolution aerial image. The resolution of the aerial image is around 4 times higher than that of range image. To register them, we manually selected 12 pair of point correspondences to find the registration transform between the range and aerial images.

3.2 Image based classification

Our goal of using aerial image is: employing texture and color information from imagery to automate the processes of extraction buildings from LiDAR, and using edges extracted from the high-resolution aerial image to refine the accuracy of LiDAR models.

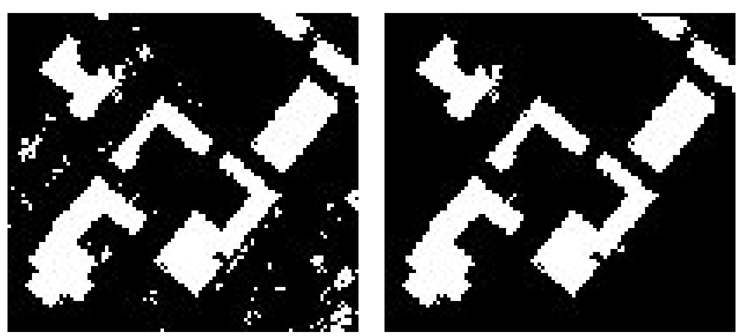

Figure 5. (a) Data classification result based on pure depth information, and (b) refined result of using color information from image.

We need to extract buildings from background, i.e. classification of the data into two sets: building set and terrain set. The range image from LiDAR contains depth for each point, so it's naturally to use it for data classification. A simple depth filter is used. Objects below a certain height threshold are classified as terrain. Otherwise, they must belong either to buildings or vegetations. Figure 5(a) shows result of using this simple depth filter to classify the range image of Figure 4. The dark parts denote terrain, while white parts are either buildings or

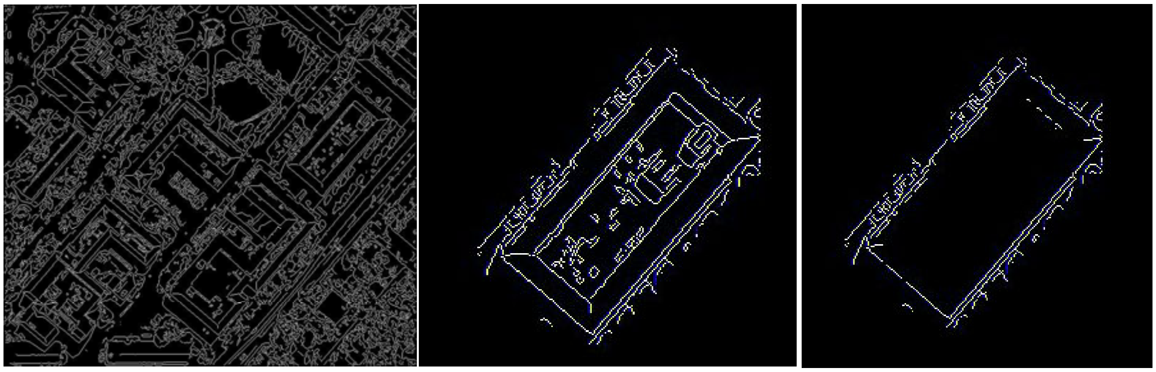

Figure 6. (a) Detected edges from aerial images, (b) refined edges near one building, and (c) refined edges near building boundary.

vegetations. However, it is difficult to further classify the vegetations from the buildings using only the height information. There is a technique pursued by using height texture to tackle this problem [Hans, 1999], however it's difficult to apply for general case, especially when the vegetations are close to buildings. Our solution to this problem is to use color information provided by the imagery source. We first map the white points (indicate building or vegetation in Figure 5) in the range image to the aerial image. We then classify those points if they belong to building or vegetation based on color information. Currently we are using a RGB color classification, i.e. if a point color is green, it is classified to vegetation set, and otherwise it belongs to building set. The refined result is shown in Figure 5 (b) in which the most vegetation areas have been correctly removed.

3.3 Edge extracting from aerial images

The Canny algorithm is used to extract building edges from aerial image. The resulted edge map is shown in Figure 6 (a). As we can see that there are too many edge points in the map that make it harder to find correct building edges. So, we use several shape cues extracted from range image to refine the result. From the segmented range image, we obtain region information of buildings. We use this region cue to filter the edge map so that only those edges near building regions are kept. Figure 6(b) shows the refined edges near a cube shape building. Still, the edge number is large that makes the number of matching hypothesis large. To further refine that, we use building boundary information obtained from the range image. We use this boundary cue to filter the edge map so that only those edges near building boundary are kept. The final result is in Figure 6(c). By using these cues from range image, we reduce dramatically the number of edges, hence the number of hypotheses of building extraction.

3.4 Hypothesis and modeling

To extract building shape, we need to group the above detected edges. General feature grouping is a classic, but still open question in computer vision research. We can however reduce the difficult by first

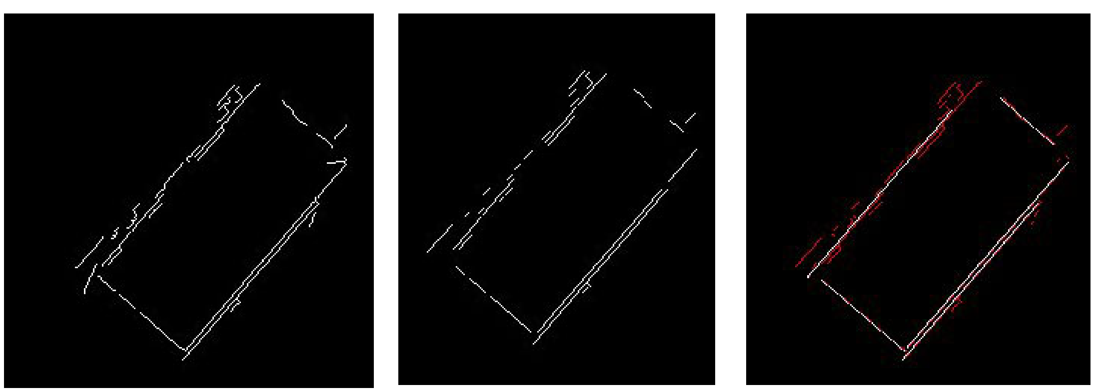

Figure 7. (a) Edges after applying a link filtering, (b) Edges after slope filter, and (c) Combined edges as hypotheses candidates.

focusing on a set of certain building shapes such as rectangle. We use a hypotheses-verification strategy to group the building edges. We first build several hypotheses for the building edges, and then verify them using the information provided from the range image. For rectangular shape, each two pair of parallel edges can form a hypothesis.

The number of possible hypotheses is quadric to the number of edges. In order to further reduce the number of hypotheses, we use several filters to reduce the number of edges. First we link the edges extracted from aerial image in anti-clockwise way. The edges with lengths less than a threshold is removed (Figure 7(a)). Second, since most of the buildings are parallel to each other. Assuming the global building direction is given, we can use a slope filter to further reduce the number of edges (Figure 7(b)). Finally, edges are combined according to geometric proximity. From the range image, we roughly know the size of building boundary. Using the building size as another length filter, we further remove the short combined edges. The result is shown in Figure 7 (c), where the red edges are edges before combination, and white edges are the resulting candidate edges for hypothesis formation.

In the hypothesis formation step, each two pair of parallel edges is selected to form one rectangular hypothesis. In Figure 7(c), for example, there are 5 edges, hence can form 5 rectangle hypotheses (Figure 8(a)). To verify the hypothesis, each rectangle is mapped back to the range image (Figure 8b). The one having maximum overlap area with the range image is selected as the building boundary. As so, we extract the building edges from aerial images for building modeling, and extract surface information from LiDAR for model fitting. The two information are then combined to obtain complete model parameters. The final modeling result is shown in Figure 8(c).

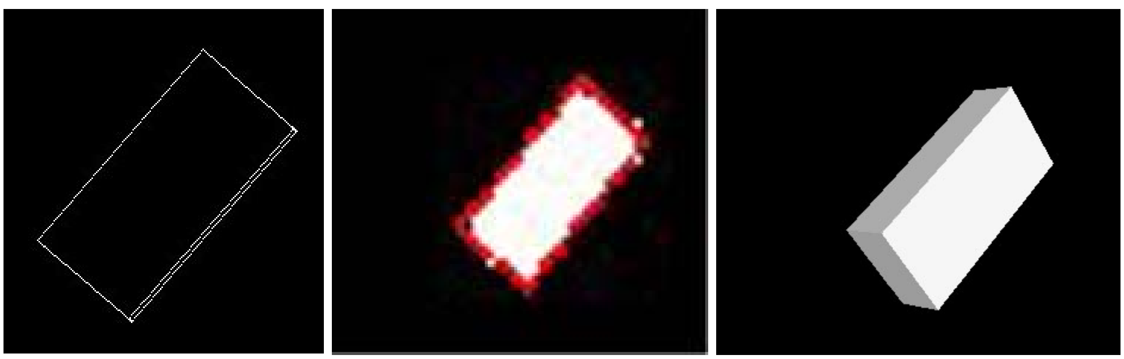

Figure 8. a) Rectangle hypotheses; b) Rectangle mapped back to range image; c) Modeling result.

3.5 Discussion

Using combined information from LiDAR data and aerial image for building modeling has several advantages: a) it improves modeling accuracy. For example, the resolution of the aerial image we used is around 4 times that of the range image. The edges extracted from aerial image are much more accurate than that detected from LiDAR data; b) it helps to automate the modeling process. For example, in our previous system, user is required to select two seed points (two user mouse clicks) to model a cube shape building. This user assistance is automated by using the above approaches; and c) it reduces computation complexity. For example, by using the shape cues from the range image we dramatically reduce the number of hypotheses (more than 90 percent), hence greatly improve the overall performance.

However, there are several issues we should pay attention to when combining information from multiple sensors. First, the different data sources should be spatially co-registered. For knowledge level information, such as cues about building regions, registration precision is not that strict. For pixel level information, such as edges and corners, then pixel level registration is required. Our work uses both knowledge level and pixel level information for modeling, so registration precision is critical. Second, images contain more information than LiDAR, such as texture and color. However, this complex information will also make the analysis complex. One problem is that there are too many edges, how to extract the correct edge information is difficult. Another problem is how to group edges to form hypothesis for complex building shapes lack of general solution. That’s why most currently automatic system can only handle some simple model shapes like rectangles.

4 Results and Conclusion

We have applied to our modeling system to a variety of dataset, and it has demonstrated its flexibility and capability for wide range of complex buildings. Figure 9 shows the results of applying the system to model a LiDAR data of Purdue campus. Due to the lack of actual measurement of the buildings, quantitative evaluation of modeling accuracy is not feasible. We use two methods for evaluation. The first method is embedding the refined model into original LiDAR model (Figure 9(c)). The second method is using imagery geo-referencing to verify the accuracy of the models (Figure 9 (d)). Both results confirm the accuracy of the proposed modeling system.

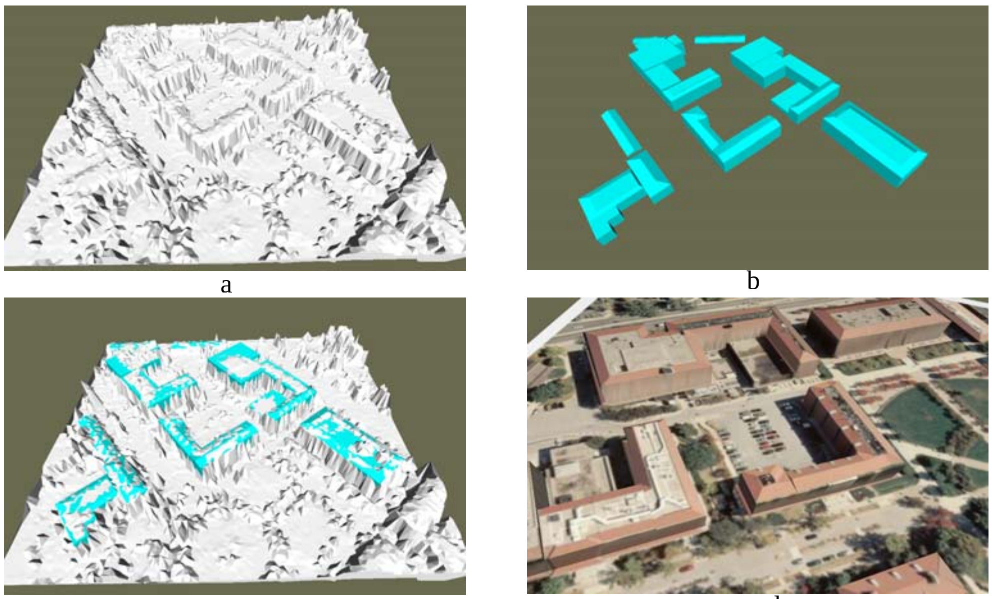

Figure 9. Complete models of Purdue dataset. (a) Original LiDAR data, (b) refined models, (c) refined models embed in original models, and (d) refined models with aerial image projection for accuracy verification.

This paper presents a modeling system using airborne LiDAR and aerial imagery. Our approach is primitive-based technique. Based on the shape of building rooftop, we classify a complex building section into two groups: linear-fitting primitives and high-order surface primitives. Once a building is segmented, geometric primitives are fit to an element's mesh model data, and the best fitting models represent the complete building structure. To improve accuracy and efficiency, we use image information to aid model and refine processes. Both the knowledge-level and pixel-level information are used. The texture and color information from aerial image is used to automate the segmentation process. Building shape cues from range image are used to reduce the number of model hypotheses and computation complexity. Edges from high resolution aerial images are used to improve the model accuracy. We demonstrated the system's flexibility and capability for modeling wide range of complex buildings.

Acknowledgement

This work was supported by the a Multidisciplinary University Research Initiative (MRUI) on "Next Generation 4-D Distributed Modeling and Visualization", and in part by the National Geospatial Intelligence Agency (NGA) under a NGA University Research Initiative (NURI) program. We thank the Integrated Media Systems Center, a National Science Foundation Engineering Research Center, for their support and facilities.

References

Ahmed F.E., James S.B. (2002). Building Extraction Using Lidar Data, *ASPRS-ACSM Annual Conference and FIG XXII Congress*.

Christian F., Z. Avideh (2001). 3D Model Generation for Cities Using Aerial Photographs and Ground Level Laser Scans. *Conference on CVPR*, Kauai, USA, pp. II-31 – 38, vol. 2.2.

George V. (2002). Fusion of Laser Scanning Data, Maps and Aerial Photographs for Building Reconstruction, *IGARSS*.

Hans G.M. (1999). Fast Determination of Parametric House Models from Dense Airborne Laser Scanner Data, *Workshop on Mobile Mapping Technology*, Bangkok, Thailand.

Hu, J., S., You, U., Neumann (2003). Approaches to Larege-Scale Urban Modeling, *IEEE Computer Graphics and Applications*, Volume 23, Issue 6, pp. 62 - 69.

Huertas, A., Z. Kim, R. Nevatia (2000), Multisensor Integration for Building Modeling, *CVPR00* (II: 203-210).

Jochen S. (2000). Combining Geometrical and Semantical Image Information for the Improvement of Digital Elevation Models, *Proceedings of the 20th EARSEL-Symposium.*

Kerry, M., K. Amnon (2002). Integration of Laser-derived DSMs and Matched Image Edges fro Generating an Accurate Surface Model, *ISPRS* 56 (167-176).

Lee, S.C., A. Huertas, and R. Nevatia (2000), Modeling 3D Complex Buildings with User Assistance. *WACV*, 2000, pp. 170-177.

Michele C., C. Bruno (1998). Optical and Radar Data Fusion for DEM Generation. *IAPRS*, Vol. 32,Part 4.

Morgan M. and K. Tempeli (2000). Automatic Building Extraction from Airborne Laser Scanning Data. *Proceeding of the 19th ISPRS Congress*, Book 3B, pp. 616-623, Amsterdam.

Neumann, U., S. You, J. Hu, B. Jiang, and J. W. Lee (2003). Augmented Virtual Environments (AVE): Dynamic Fusion of Imagery and 3D Models, *IEEE Virtual Reality 2003*, pp. 61-67, Los Angeles California.

Norbert H. and B. Claus (1997). Generation of 3D city models from Airborne Laser Scanning Data, *3rd EARSEL Workshop on LIDAR remote sensing of land and sea*.

Ribarsky, W., T. Wasilewski, and N. Faust (2002). From Urban Terrain Models to Visible Cities, *IEEE Computer Graphics & Applications*, Vol. 22.

Robin D. M., V. T. Udo and C. Peter (2001). High Resolution Surface Geometry and Albedo by Combing Laser Altimetry and Visible Images, *ISPRS*.

Seresht M. and A. Azizi (2000). Automatic Building Recognition from Digital Aerial Images. *Proceeding of the 19th ISPRS Congress*, Book 3B, pages : 792-798, Amsterdam.

You, S., J. Hu, U. Neumann, and P. Fox (2003). Urban Site Modeling From LiDAR, *Second International Workshop on Computer Graphics and Geometric Modeling CGGM'2003*, Montreal, CANADA.

Zhao, B., and J. Trinder (2000). Integrated Approach Based Automatic Building Extraction. *19th ISPRS Congress*, Book 3B, pages : 1026-1032, Amsterdam.