EXTRACTION OF BUILDINGS FROM SATELLITE IMAGE AND LASER RANGING DATA

TAO GUO AND YOSHIFUMI YASUOKA
Institute of Industrial Science, the University of Tokyo
4-6-1 Komaba, Meguro, Tokyo 153-8505, Japan

ABSTRACT

Within the paper an approach for the automatic extraction and reconstruction of buildings in urban built-up areas base on fusion of high-resolution satellite image and laser ranging data is presented. The presented data fusion scheme is essentially motivated by the fact that image and range data are quite complementary for the purpose of 3D feature extraction. Raised urban objects are first segmented from the terrain surface in the laser ranging data by making use of the spectral signature derived from satellite image, afterwards building potential regions are initially detected in a hierarchical scheme. A novel 3D building reconstruction model is also presented based on the assumption that most buildings can be approximately decomposed into rectilinear polyhedral patches. With the constraints of presented building model, building linear features are used to generate the hypothesis, and then the verification processes follow. And a subsequent logical processing of the primitive geometric patches leads to 3D reconstruction of buildings with good details of shape. The approach is applied on the test sites and shows a good performance, an evaluation is described as well in the paper.

1. INTRODUCTION

Modeling urban entities in geometric forms provides a fundamental platform for urban related studies. Traditional 2D map has been long one of popular and valuable urban models, but there are several inherent constraints which limit itself from widely using in this field, among them: it takes much resources and a long cycle for updating, field survey and aerial photography measurement are quite expensive and time consuming, meanwhile, the change in urban areas normally occurs very fast and frequently, for some near real time applications, updating changes in short cycle is normally required; Moreover 3D information can be not properly presented, even though the information of the third dimension could annotated on 2D maps, the truly and fully 3D representations of real-world objects are still not available. Moreover 3D information is quite significant for many applications such as urban safety analysis, urban microclimate and pollution control analysis, transportation navigation, landscape planning and visualization, telecommunication industry etc.

In comparison with field survey, deriving 2D or even 3D information from the photography is quite economic and fast way. Meanwhile remote sensing generally is the science of acquiring and analyzing information about objects or phenomena from a distance. From the early use of aerial photography, up to the later use of satellite imagery, remote sensing has been recognized as a quite valuable technology for environment study.

In last decade, the resolution of satellite imagery has reached from several kilometers to several dozen meters, and shown a really good performance in the study of large homogeneous area, i.e. land use classification, and geological structure derivation etc. Still with this level resolution, it is not sufficient yet for the local scale studies, which require relatively detailed information. However recently when more and more high-resolution satellite imageries have become widely available, the application of remote sensing technology for the acquisition of the details of urban, not only land use classification, is currently becoming an issue of increasing interest and high importance to both researchers and users from remote sensing, geo-information, urban planning, government administration, commercial companies, etc. This field is often referred as urban remote sensing as well.

Despite the optimistic potentiality it is a challenging yet extremely difficult task to automatically even semi-automatically extract objects in urban densely built-up environment solely based on image, this is not only because of the attributes of image, but also because of the high object density, occlusions and scene complexity in urban environment. In this research, we combine the high-resolution IKONOS satellite images and airborne laser scanning data to extract buildings for the purpose of 3D city modeling based on such a fact that the characteristics of these two kinds of data source are quite complementary for 3D features extraction.

2. BUILDING SEGMENTATION

2.1 Urban Terrain Surface Detection

Segmentation is the most common step involving object detection, the goal of segmentation normally is not to locate object, but to partition data sources into target region. In particular, for city modeling purpose, due to the extremely complex scene, feasible simplification of urban objects becomes very significant for late processing.

In our scheme, because the building is the focus of interests, the purpose of the first step is trying to separate raised objects from terrain, and to make a rough classification on raised objects. The task sounds simple, but is actually really tough and the results have significant influence on late processing. So far accurate urban terrain automatic detection still remains as a challenging topic. Our scheme starts with a rough classification by using spectral signature such as well-known NDVI derived from IKONOS multiple-spectral image to exclude most trees, and then height information taken from Digital Surface Model (DSM), which is directly generated from airborne laser scanning data, is adopted to help for finding pixels on the ground. A ground trend surface can be constructed if sufficient ground pixels are found. By comparing with the ground trend surface, tall and raised objects which violate the ground trend very much are therefore excluded. Up to this stage, the objects with the high possibility of being non-terrain features have been "picked out", and then terrain surface can be detected with better accuracy by using well-recognized morphological operations, which is suffering with the difficulty of elimination of big size objects. Subtracting the terrain surface from the original DSM results in a representation of most objects rising from the terrain approximately put on a plane, and the slope influence of terrain has been removed, that is so-called normalized DSM (nDSM).

2.2 Buildings Potential Regions Detection

Generation of normalized DSM leads to a coarse classification on urban objects, and next step towards extraction of 3D building is to locate the building potential regions. This step can be seen as the further refinement of urban classification in terms of both class and accuracy. Solely based on individual data source is still not sufficient for this purpose due to the same reasons as described above. The practical solution is to derive complementary and reliable clues and carefully integrate them in a reliable way.

In order to combine IKONOS image and nDSM, they have to be co-registered. Because the ortho-image is rectified by using ground control points, this means objects have been ortho-rectified only at their footprint parts, tall objects appear still lean on the image, here is referred as "building lean" problem. The co-registration of image and nDSM becomes an issue of "restore lean buildings to straight". Based on the attributes and geometric relations of both IKONOS image and ALS data, buildings are then ortho-rectified using nDSM. But practical issues still need to be considered: due to the existence of noise and displacement in DSM data, output image will appear noisy, so it is suggested that DSM data is smoothed before actually use. Second, some information-loss patches will occur in output image, though appearance of image can be improved through applying interpolation and re-sampling, these patches degrade the image quality in a certain degree. Anyway, while solving building lean problem, the quality of output image, especially planar accuracy is distorted at some extent. Thus a common goal of practical application is to seek a trade-off. Because when elevation of object becomes low, this shift becomes small as well. Within some range, this shift can be omitted. An elevation threshold is suggested to set, and only tall objects above this threshold are ortho-rectified. In most cases, this solution produces a reasonable output. After solving the building lean problem, the IKONOS image and normalized DSM have been actually co-registered.

Multiple clues derived from co-registered IKONOS and nDSM are ready for combination so far. A hierarchical scheme for detection of building potential regions from above coarse classification is presented as illustrated in Figure 1. The essential idea is to utilize complementary attributes of multiple clues for separating objects by emphasizing various characteristics at different stages. In our schemes, clues of height, size, spectral signature, and texture are hierarchically used as discrimination factors.

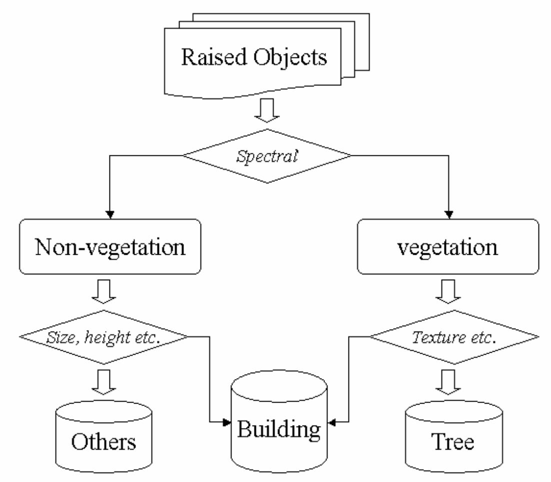

Figure 1: General scheme for building detection

At this stage, buildings are detected as 2D regions, but no precise descriptions of 2D buildings are given, conversely, a "region relaxation" is made, this because detailed construction of 3D buildings will be conducted in the next step, description of 2D buildings naturally becomes available once 3D buildings geometric model are reconstructed. Fusion of image and height data plays a very important role in this processing.

3. BUILDING EXTRACTION

3.1 Building Model

Building detection has given an explicit representation for buildings in terms of region and location. The goal of building extraction is to generate a detailed logical description about the building in 3D space. In order to make it sense what kind of 3D building model and how detail-level can be expected from the specific data sources, firstly it is necessary to give an explicit semantic description on buildings, which is so called "building semantic model" in this paper. The idea behind is that instead of trying to directly extract building, firstly go through the data sources and investigate what the building appears in these data sets, afterwards, a specific 3D building model can be determined from such a semantic description about building appearance, thus derivation of features becomes a model-orientated process for generating reliable evidences. Therefore this strategy leads to a better simplicity and efficiency.

Our building semantic model considers several factors of buildings: size, shape, planes of surface, and structure. Therefore, for IKONOS image and ALS data, the building semantic descriptions is given as:

- With sufficient size for reliable recognition.
  Small size buildings could be missed or merged to big ones with neighbors.
- Rectilinear outline.
  Neighboring edges of buildings are straight lines and form orthogonal polygons, which could be concave and have holes.
- With multiple layers of roof patches structure.
  Information about walls is unable to be directly extracted. Complex structure of buildings is represented as composition of roof layers.

Correspondingly, "buildings parameterized model" as (equation 1) is defined according to building semantic model, and it is specially the geometric description about building. According to the buildings parameterized model, 3D roof patches reconstruction becomes the key for 3D building reconstruction.

$$\begin{aligned} &3DBuilding = f_{\mathrm{mod}el}(Location(Center(x, y), Size), Orientation, \\ &\cup f_{layer}(Height, f_{roof\_patch}((x, y)_n))) \end{aligned} \tag{1}$$

The presented 3D building model as illustrated in Figure 2, can be considered as a constructive solid geometry (CSG) model, where the primitive actually is the rectilinear polyhedron.

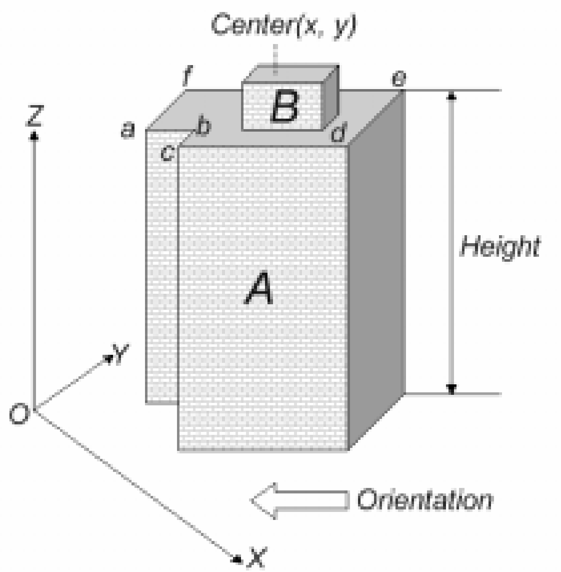

Figure 2: Building parameterized model

3.2 3D Building Reconstruction

After clarify the building model based on the input data, our scheme starts with feature extraction, here interest is feature of lines. Edge detectors are used to detect edges from both image and nDSM in the detected building potential regions, and then straight lines are extracted by using Douglas straight-line fitting algorithm (Douglas and Peucker, 1973). Extracted straight lines could be the boundaries of buildings, either could be edges of other objects such as noise or boundaries of shadows etc. Our scheme benefits very much from the geometric constraints of the buildings rectilinear model, that is, if a straight-line is a boundary of building, its direction must be parallel or perpendicular to the building orientation. Thus the task is greatly simplified and becomes an issue of detection of building orientation.

A method based on statistic analysis of straight lines direction histogram to determine building orientation is presented. And then straight lines are grouped into primary and secondary clusters according to their directions. If a straight line, within a predefine range is parallel to the building orientation, it is grouped into primary cluster, if it is perpendicular to building orientation, then put it into secondary cluster. Others are put into abandoned cluster.

The scheme of 3D roof patches reconstruction as illustrated in Figure 3 is typically formulated as a process of hypotheses generation and verification. In order to generate 3D hypotheses, the first step is to transform necessary features from 2D into 3D space by constructing 3D layers. Afterwards, on each layer, the hypotheses are initialized as rectangles covering the building potential region. Then the rectangle is further decomposed into small rectangles using the clues of primary and secondary straight lines. Verification is performed to each rectangle, and finally 3D roof patches are reconstructed by merging valid rectangles. 3D roof patches and their height form actually orthogonal polyhedrons, with these polyhedrons, the 3D building reconstruction becomes a simple issue of logical union operation applying to 3D building primitives.

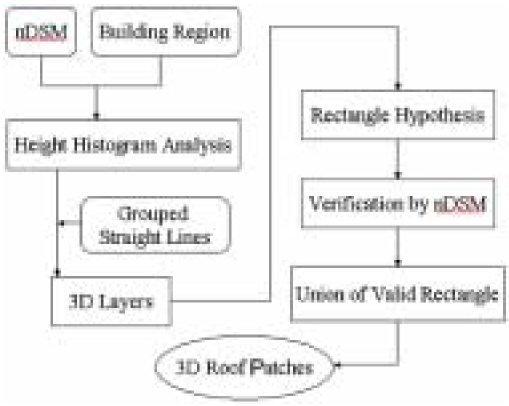

Figure 3: 3D roof patches reconstruction scheme

4. RESULTS AND VALIDATION

Presented schemes have been conducted in two test sites, one is around the area of the Komaba campus of the University of Tokyo, and another is an area in Shinjuku of Tokyo. Meanwhile a qualitative and quantitative assessment about the presented 3D building extraction model is briefly conducted based on these experiments. For 2D assessment, by comparing with digital maps with our model, it shows that most buildings (about 85-87%) in urban densely built-up areas can be approximately extracted by applying our 3D rectilinear building model from IKONOS image and laser scanning data. And for modeling rectilinear building, our 3D building extraction scheme shows a very good accuracy. On horizontal ground plane, planar accuracy is within one meter. For the comprehensive comparison of extracted 3D building models is actually unable to conduct due to the unavailability of 3D reference data. In order to give a primarily qualitative 3D evaluation, we collected the architecture design plan of our institute building and manually compare with our generated model as shown in Figure 4, the results show our scheme is able to achieve within one-meter accuracy in elevation and the performance is quite promising. Also a simple visualization of our 3D building models is shown in Figure 5.

Figure 4: Reconstructed 3D building model.

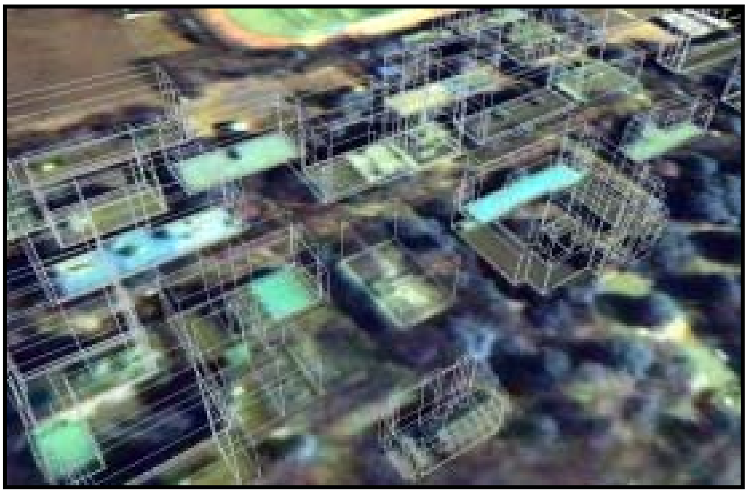

Figure 5: Simple visualization of 3D building models

5. CONCLUSIONS

A novel 3D building extraction scheme has been presented in this paper. It consists of generally three steps: urban objects segmentation, building detection, and 3D building reconstruction. In each step, image and laser ranging data are intensively fused at different level. The scheme starts with the low-level pixel processing for feature segmentation and detection, and then exploits available knowledge about building appearances in the data sources to group detected features into organized primitive elements, finally extract 3D geometric description of building by using a recursive process of hypotheses generation and verification. And also conducted experiments demonstrate a promising performance for presented schemes. Future work will be directed towards improving each individual component of the presented 3D building extraction schemes, trying to generalize our models, and the attempt of applying our models into practical applications is a significant aspect as well.

ACKNOWLEDGEMENT

This work is carried out as a part of the joint research of Japan Space Imaging Corporation (JSI) and the Institute of Industrial Science (IIS) of the University of Tokyo. We would like to thank JSI for providing us IKONOS images and ALS data.

REFERENCES

Douglas, D. H., and Peucker, T. K., 1973. Algorithms for the Reduction of the Number of Points Required to Represent a Digitized Line or Its Caricature. *Canadian Cartographer* 10(2), 112-122.

Forstner, W., 1999. 3D-City Models: Automatic and Semiautomatic Acquisition Methods. In D.Fritsch and R.Spiller (eds.): *Photogrammetric Week'99*, Wichmann, Karlsruhe.

Fraser, C. S., Baltsavias E., and Gruen, A., 2001. 3D Building Reconstruction from High-resolution IKONOS Stereo-imagery. In Baltsavias et al. (eds), *Automatic Extraction of Man-made Objects from Aerial and Space Image (III)*, 331-344.

Guo, T., 2003. 3D City Modeling Using High-resolution Satellite Image and Airborne Laser Scanning Data. Doctoral dissertation of the University of Tokyo.

Guo, T., and Yasuoka, Y., 2002. Data Fusion of High-resolution Satellite Images and Airborne Laser Scanning Data for Building Detection in Urban Environment. *SPIE Third International Asia-Pacific Environmental Remote Sensing Symposium, 2002.*

Haala, N., and Brenner, C., 1999. Extraction of Buildings and Trees in Urban Environments. *ISPRS Journal of Photogrammetry & Remote Sensing*, 54, 130-137.