Paper presented at the 22nd Asian Conference on Remote Sensing, 5 - 9 November 2001, Singapore.

EXPERIMENTS ON CSG MODEL-BASED BUILDING EXTRACTION FROM AERIAL IMAGES

Yi-Hsing TSENG[1] Sendo WANG[2]

Department of Surveying Engineering, National Cheng Kung University
#1 University Road, Tainan 701
Taiwan, R.O.C.

KEY WORDS: Building Extraction, Building Reconstruction, Constructive Solid Geometry, Model-Image Fitting, Semi-Automatic, Digital Photogrammetry.

ABSTRACT: Building extraction based on pre-established models has been considered a promising approach to acquiring precise, reliable and complete 3D data of buildings from aerial images. Based on the previously proposed CSG model-based building extraction, complex buildings are modeled as a combination of volumetric primitives by the theory of Constructive Solid Geometry (CSG). The concept of the approach is to let high-level tasks (building detection, model selection, and attribution) be carried out interactively by the operator and optimal model-image fitting be performed automatically by a computer algorithm. The shape and pose parameters associated to a primitive provide a link between perception (images) and prior knowledge (primitive) of a building part, so that the fitting method proceeds to determine the shape and pose parameters so as to fit a primitive with the corresponding images. Having all of the building parts been uniquely represented by parametric primitives, a building can be reconstructed by using CSG Boolean set operators to combine the building parts. Consequently, a building is represented by a CSG-tree in which each node links two branches of combined parts. This paper demonstrates 10 examples of extracting various buildings. All of the tests were performed in the prototypal system implemented in a CAD-based environment cooperated with a number of specially designed programs. The process time for each primitive is about 20 sec and the successful rate of model-image fitting is about 90%. The test results are encouraging and revealing the directions of improving the CSG model-based building extraction.

1. INTRODUCTION

Acquisition of 3D data of city objects has become a topic of increasing importance. Intensive research activities have been engaged in both the photogrammetry and the computer vision communities (Braun et al., 1995; Grün, 2000; Lang and Förstner, 1996; Mohan and Nevatia, 1989; Shufelt, 1999; van den Heuvel, 2000; Vosselman and Veldhuis, 1999; Weidner and Förstner, 1995) aiming for automatic or semi-automatic building extraction from digital aerial images. Building extraction based on pre-established models has been considered a promising approach to acquiring precise, reliable and complete 3D data of buildings from aerial images. We previously proposed a semi-automatic approach to model-based building extraction (Tseng & Wang, 2001). In this approach, complex buildings are modeled as a combination of volumetric primitives by the theory of Constructive Solid Geometry (CSG). The concept of the approach is to let high-level tasks (building detection, model selection, and attribution) be carried out interactively by the operator and optimal model-image fitting be performed automatically by a computer algorithm. The shape and pose parameters associated to a primitive provide a link between perception (images) and prior knowledge (primitive) of a building part, so that the fitting method proceeds to determine the shape and pose parameters so as to fit a primitive with the corresponding images. Having all of the building parts been uniquely represented by parametric primitives, a building can be reconstructed by using CSG Boolean set operators to combine the building parts. Consequently, a building is represented by a CSG-tree in which each node links two branches of combined parts. This approach is developed with the prospects of releasing the operator from tedious point measurement and efficiently delivering precise and reliable results. In contrast with the traditional point-by-point digitization mapping process, this approach promotes an object-by-object data acquisition procedure.

In this paper, we firstly made a brief introduction to the strategies and workflow of the CSG model-based building extraction. And then, the concepts of building modeling, model-image fitting, model combination and constraints are also briefly discussed. Finally, we focus our view on the investigation into a series of experimental tests in the aspects of model availability, working efficiency, needed constraints, success rate of model-image fitting, and fitting accuracy etc. Ten various buildings were arbitrarily selected for the test. The experimental results are demonstrated and analyzed to evaluate the performance of the proposed algorithm. The prospect of this study is to reveal the directions of improving the CSG model-based building extraction.

---

[1] Associate Professor, Phone: +886-6-2757575-63835, Fax: +886-6-2375764, E-mail: tseng@mail.ncku.edu.tw
[2] Ph.D. Student, Phone: +886-6-2373876-834 Fax: +886-6-2375764, E-mail: sendo@sv.ncku.edu.tw

2. CSG MODEL-BASED BUILDING EXTRACTION

2.1 Strategies and Workflow

Based on the CSG principle, buildings are modeled as a combination of volumetric primitives. A primitive may represent a building or a part of building depending on the complexity of building shapes. Buildings with complex topology can be reconstructed by Boolean operations of the building part primitives in a generic way. The system should provide a **model base** in which various parametric primitives are included. Each model is associated with a number of **shape parameters** and **pose parameters**. The representation of a building part is implemented by setting the values of shape and pose parameters for the representative model. The operator needs to find an appropriate model from the model base corresponding to the target (building part). Some interactive user interfaces should be provided for the operator to select a model and to perform an approximate fitting between model and images. Similar to the traditional procedure of photogrammetric mapping, building extraction is working based on the assumption that the interior and exterior orientation data of the images are known, so that the fitting can be examined by projecting an instantiated primitive onto the images. Then, optimal model-image fitting is performed by the system for each primitive. Both the manual and automatic fitting processes are achieved by adjusting the shape and pose parameters of a model to fit the target images. Having determined all of the building parts, the whole building can be reconstructed by combining the building parts with generic Boolean operations. It needs to introduce some constraints and local modifications of primitives during the process of combination. Again, interactive user interfaces should be provided for the operator to specify various constraints and modifications. The results of this building extraction approach will be a CSG-tree in which each node links two branches of combined building parts. The leaves of the tree are building primitives and each of them is associated with the shape and pose parameters.

The workflow of this approach includes four stages: model selection, approximate fitting, optimal fitting, and primitive combination (Figure 1). In the first stage, the operator detects buildings in a navigation mode. The operator then needs to analyze the building and to divide the building into parts that can be modeled by the primitives predefined in the model base. The second stage is an interactive procedure to revise the shape and pose parameters of a model to fit the interested building part approximately. This procedure can be done by providing the user a dialog window to adjust the shape and pose parameters and showing the model wire frame on the images for checking. The third stage is an automatic fitting procedure. Started from the approximate fitting, the optimal fitting is achieved iteratively by using the least-squares model-image fitting algorithm. The final stage again is an interactive procedure to combine building primitives using Boolean operators. Some attachment constraints and local modifications can be specified for the combination process. Through the workflow, the system does not require stereo viewing and point measurements.

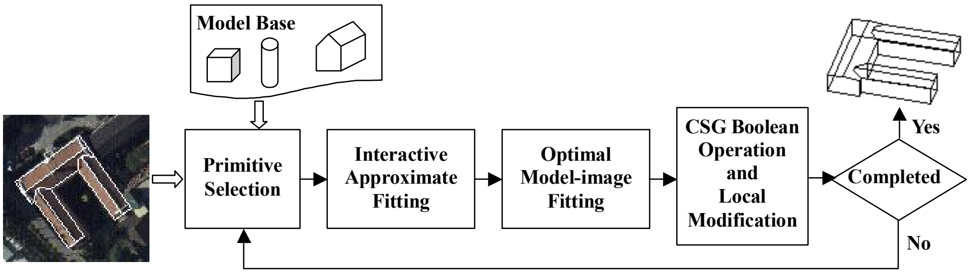


Figure 1. The workflow of the CSG model-based building extraction.

2.2 Building Modeling

Buildings are modeled as a combination of volumetric primitives. The parameters associated with a primitive can be categorized into shape and pose parameters. The parametric changes would not affect the intrinsic geometric properties. For example, a solid-box primitive is able to represent a rectangular building (or building part) with the shape parameters of length (l), width (w), and height (h). By changing the shape parameters, the primitive can be scaled or elongated in each dimension to fit the size of a rectangular building. Different primitive models will be associated with different shape parameters. For example, a gable-roof primitive model should have an additional parameter to model its roof height (rh). Unlike the from shape parameters, pose parameters are not associated to the changes in size or shape, but define the position and orientation of primitives. In a three-dimensional space, it is adequate to use 3 translation parameters (dX, dY, dZ) and 3 rotation parameters, tilt, swing, and azimuth (t, s, α), to depict the position and orientation of an object. However, most buildings should be kept vertical, so that the tilt and swing parameters can be turned off. Therefore, one can use 4 pose parameters (dX, dY, dZ, α) for all kinds of building primitives (Suveg and Vosselman, 2000 and Vosselman and Veldhuis, 1999).

The coordinate systems involved in this approach include **model**, **object**, **photo**, and **image** coordinate systems. Transformations between coordinate systems can be performed based on associated parameters (Figure 2). A model is defined in the model coordinate system, and can be transformed into the object space in accordance with the shape and pose parameters to represent a building part. Through a central projection, a building model can be transformed into a 2D photo coordinate system in accordance with the known exterior orientation. Furthermore, photo coordinates can be transformed into the image coordinate system in accordance with the interior orientation.

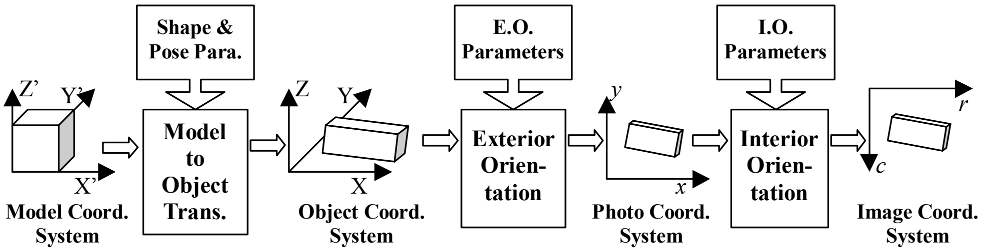


Figure 2: The coordinate systems involved in this approach and their relationships.

2.3 Model-image Fitting

The principle of model-image fitting is to adjust shape and pose parameters so as to fit model with the corresponding features extracted from the images. Features for matching are edge lines, so that the best fit is achieved by minimizing the sum of the perpendicular distances from the edge pixels to the projected edge line. The Least-squares Model-image Fitting (LSMIF) originally proposed by Lowe (1991) is modified to solve for projection and model parameters. This fitting algorithm starts with an approximate fit, and then iteratively converges to the optimal fit.

The measure equation models the perpendicular distance from an edge pixel to the projected edge line. An edge line of a primitive model defined in the model space should be transformed into the object space using the shape and pose parameters and then transformed into the photo coordinate system using the exterior parameters. On the other hand, an edge pixel derived in the image space should also be transformed into the photo coordinate system using the interior parameters. Denoting the photo coordinates of the two end points of the edge line as (x1, y1) and (x2, y2) and the photo coordinates of an edge pixel as (x, y), the equation will be:

$$v = \frac{(y_1 - y_2)x + (x_2 - x_1)y + (y_2 x_1 - y_1 x_2)}{\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}}$$

in which, *v* is the perpendicular distance from the edge pixel to the edge line. Coordinate x1, y1, x2, and y2 are functions of the shape and pose parameters as well as exterior orientation parameters. Because the exterior orientation is known, the unknowns of the equation are the shape and pose parameters. For the case of box primitive, the measure equation can be denoted as: $v = F(l, w, h, dX, dY, dZ, \alpha)$. This equation is a non-linear function with respect to the unknowns. An edge pixel within the buffer of an edge line in each photo will form a measure equation. The number of equations will be much more than the number of unknowns, so that it is an over-determining system. By using Newton's method, the iterative least-squares approach is a standard method to pursue the optimal solution.

2.4 Boolean Set Operations and Local Modification

In the system, building parts are determined one by one, which should be combined to form a complete building using Boolean set operations, such as union ($\bigcup$), intersection ($\bigcap$), and difference (–). Attachment of building parts is a case of union that needs special care. An attachment operation usually should incorporate some constraints and local modifications to ensure that a connection between two primitives makes sense. However, it is inevitable to have a discrepancy or overlap in-between the two primitives because of fitting errors or uncertainties. Consequently, local modifications on one primitive or on both primitives are required if an attachment constraint is incorporated in the union of these two primitives. Again, in order to maintain the intrinsic geometric properties, all local modifications should be carried out by changing shape or pose parameters.

Attachment constraints for combining building primitives can be categorized into facet-to-facet, edge-to-edge, and orientation alignment constraints. A facet-to-facet constraint is needed for side-by-side connections of two building primitives, such as the examples in Figure 3(a). When two primitives need to connect side by side and some edges of the connected facets need to overlap, edge-to-edge constraints should be applied. Figures 3(b) and (c) are examples requiring edge-to-edge constraints. Orientation alignment is also frequently required to obtain a reasonable connection. For example, a building with the structure of stacking-up boxes inherits orientation alignment among the boxes as shown in Figure 3(d).

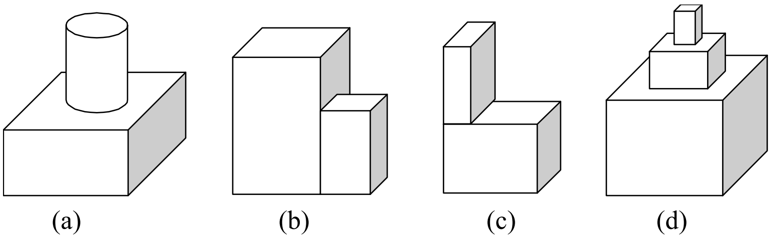


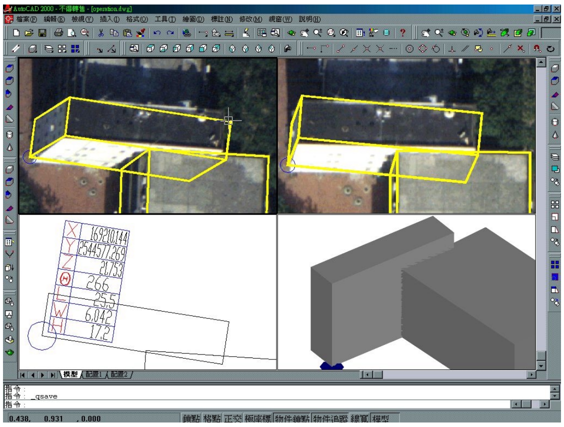
Figure 3: Examples of combining primitives with various constraints.

3. EXPERIMENTS

3.1 Implementation

The proposed approach is implemented in a CAD-based environment cooperated with a number of specially designed programs. The model base is established in the Auto-CAD system. The methods for model selection, approximate fitting, and visualization are implanted in the CAD system using Visual Basic for Application (VBA) programming. The least-squares model-image fitting is currently an independent function developed using C code. However, it would be possible to combine all of the processes into a single system.

Figure 4 shows the designed working environment. The graphical user interface allows the operator to zoom and view overlapped images in the image windows, and pick a suitable primitive from the model icons listed in the left column. Therefore, the operator can perform model selection and approximate fitting in this environment and visually supervise the fitting procedure. Each primitive is associated with an anchor point which is a point locates on the origin of the model coordinate system. By specifying the projected positions of the anchor point on the images, the approximate location of the model in the object space can be determined. Furthermore, one can obtain the approximate scale and orientation of the model by specifying the two neighbor corners of the anchor point in one image. On the screen, the lower-left and lower-right windows show the top and perspective views of the model, respectively, in the object space. The model wire frame is also superimposed on the images. After approximate fitting has been carried out, the system then iteratively solves for the shape and pose parameters to possess the best fit between model and images. Figure 5(a) shows the initial state of fitting, and Figure 5(b) demonstrates the solved final fitting.



X 169210.144
Y 2544577.269
Z 21.753
Θ 266
L 25.5
W 6.042
H 17.2

```
指令 :
指令 : _qsave
指令 :
```

0.438, 0.931 , 0.000

Figure 4: The CAD-based working environment.

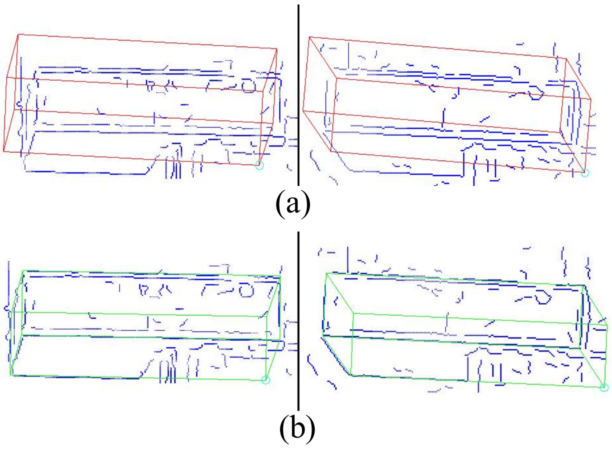


Figure 5: The (a)initial and (b)final model-image fitting.

3.2 Tests and Analysis

The test data are digitized aerial photos of the NCKU campus. The photo end-lap is 60% and side-lap is 30%. In this test, buildings were extracted from the stereo image pairs formed by end-lap. Ten various buildings were arbitrarily selected as examples 1~10 for the test. Figure 6 graphically shows the results of the 10 examples. In the aspect of model availability, all of the buildings can be properly represented by a combination of box and gable-roof primitives. Except the first example, the other buildings have to be modeled using two or more than two primitives. It can be seen that the CSG modeling is very adaptive to complex buildings. The proposed model-image fitting function is quite efficient. For each primitive, it takes about 20 sec to go through the procedure. It is much faster than point-by-point manual measurement. The successful rate of optimal model-image fitting using LSMIF is about 90% (24 out of 27 primitives). By introducing proper constraints, all cases can be solved and proper topology between connected primitives can be maintained. From the extracted spatial information of buildings, the coordinates of building corners were derived to compare with manual measurements of the corresponding points. Table 1 shows the average and RMS differences of the coordinates. Large differences appeared mostly due to confusing edges or self-occlusion. This empirical accuracy analysis evidences that the new approach generates qualified data for 3D spatial information systems.

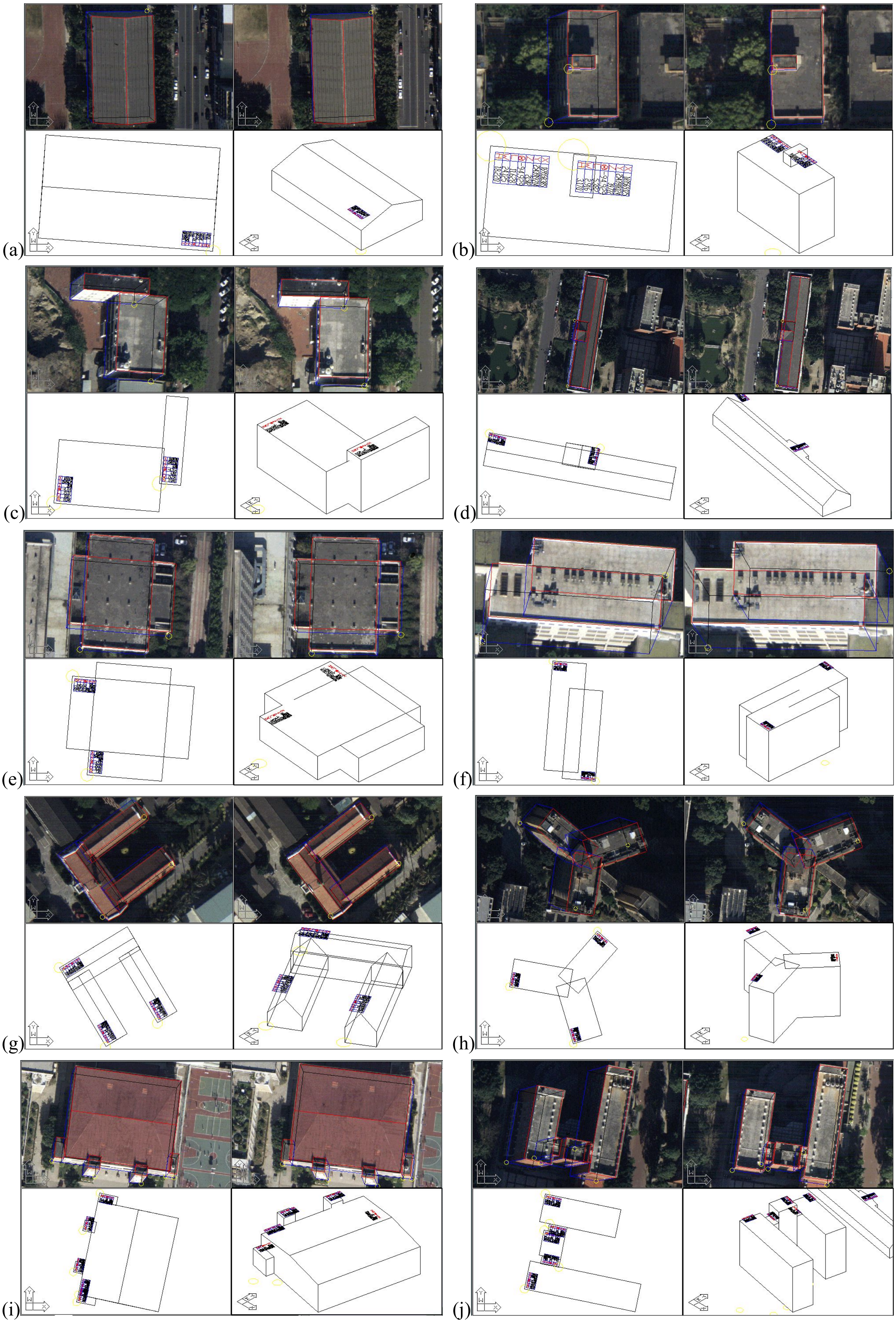


Figure 6: From (a) to (j) sequentially show the 10 test results. For each example, the stereo images are shown with superimposed model wire frames in the upper, and the top and perspective views are shown in the lower.

Table 1: The average and RMS differences of the building-corner coordinates derived from building extraction method and manual measurement.

	X (m)	Y (m)	Z (m)
Average Diff.	0.161	0.007	0.047
RMS Diff.	0.330	0.277	1.034

4. CONCLUSIONS

Based on a series of experimental tests, CSG model-based building extraction from aerial images has proved itself capable of acquiring precise, reliable and complete 3D data of buildings. This approach is much more efficient than point-by-point manual building extraction. Besides, there are some advantages of using CSG model-based building extraction. First, it employs CSG modeling so that complex buildings can be modeled with a small set of primitives. Second, the use of building model is essential while the projection of 3D objects into 2D images leads to a loss of relevant information for building extraction or useful information in the images is confused with irrelevant information. Furthermore, The final products are CSG building representations, which have the prospects of proving the fundamental data for a 3D city spatial information system.

The developed semi-automated procedures combine the human ability of image understanding with the number-crunching capacity of computers. The following aspects characterize the procedures:

- This approach lets the operator deal with high-level tasks interactively and performs optimal model-image fitting automatically.
- It releases the operator from tedious point measurement and efficiently delivers precise and reliable results. In contrast with the traditional point-by-point digitization mapping process, this approach promotes an object-by-object data acquisition procedure.
- It does not require stereo viewing or measurement.

REFERENCES

Braun, C., T. H. Kolbe, F. Lang, W. Schickler, V. Steinhage, A. B. Cremers, W. Förstner and L. Plümer, 1995. Models for Photogrammetric Building Reconstruction, *Computers & Graphics*, 19(1), pp. 109-118.

Grün, A., 2000. Semi-automated Approaches to Site Recording and Modeling, *International Archives of Photogrammetry and Remote Sensing*, Amsterdam, XXXIII, pp. 309-318.

Lang, F. and W. Förstner, 1996. 3D-City Modeling with a Digital One-eye Stereo System, *International Archives of Photogrammetry and Remote Sensing*, Vienna, 31, pp. 415-420.

Lowe, D. G., 1991. Fitting Parameterized Three-dimensional Models to Images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(5), pp. 441-450.

Mohan, R. and R. Nevatia, 1989. Using Perceptual Organization to Extract 3-D Structures, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(11), pp. 1121-1139.

Shufelt, J. A., 1999. Performance Evaluation and Analysis of Monocular Building Extraction From Aerial Imagery, *IEEE Transactions on Geoscience and Remote Sensing*, 21(4), pp. 311-326.

Suveg, I. and G. Vosselman, 2000. 3D Reconstruction of Building Models, *International Archives of Photogrammetry and Remote Sensing*, XXXIII, Part B2, pp. 538-545.

Tseng, Y.H. and S. Wang, 2001. Semi-automatic Building Extraction by Matching CSG Primitives with Aerial Images, *Proceedings of Asia GIS 2001*, Tokyo.

van den Heuvel, F. A., 2000. Trends in CAD-Based Photogrammetric Measurement, *International Archives of Photogrammetry and Remote Sensing*, Amsterdam, XXXIII, Part B5, pp. 852-863.

Vosselman, G. and H. Veldhuis, 1999. Mapping by Dragging and Fitting of Wire-Frame Models, *Photogrammetric Engineering & Remote Sensing*, 65(7), pp. 769-776.

Weidner, U. and W. Förstner, 1995. Towards Automatic Building Extraction from High Resolution Digital Elevation Models, *ISPRS Journal of Photogrammetry and Remote Sensing*, 50(4), pp. 38-49.