Proceedings of the 6th AGILE
April 24th-26th, 2003 – Lyon, France

MULTI-VIEW 3D CITY MODEL GENERATION WITH IMAGE SEQUENCES

Yongjun Zhang   Zuxun Zhang   Jianqing Zhang

School of Remote Sensing and Information Engineering, Wuhan Univ., 430079, P.R. China

1. INTRODUCTION

Three-dimensional (3D) city models are usually comprised of a description of terrain, streets, buildings, vegetations, and other man-made objects. In principle, 3D graphics has many advantages over video or 2D graphics. 3D interactive environments are an intuitive and user-friendly way to view location-based information. For many applications, 3D representation of city models improves the usability of data. People are becoming increasingly conscious about their technical, environmental and commercial possibilities (Forstner, 1999, Brenner, 1999), such as car navigation and service browsing, tourism and marketing, architecture and town planning, city climate, noise propagation and environmental research, and new landmark-based navigation systems.

Interests in 3D city model generation and quickly updating have raised significantly in the past years. In “Real Time” GIS areas, 3D city model generation and updating speed is of most importance. It became clear early that photogrammetry is able to provide a means to collect the required three-dimensional information. However, standard photogrammetric systems are not the most efficient solution to measure city models, since they were made to measure points rather than structured objects mainly composed of line segments.

This paper will discuss several important techniques toward quickly coarse-to-fine 3D city model generation. Automatic aerial triangulation and transformation are descripted to give initial values of camera parameters of each image. As for building modelling, a multi-view interactive coarse-to-fine reconstruction algorithm with low-accuracy 2D ground plan is addressed. Finally, rectification and automatic noise removing methods are adopted to generate realistic textures of buildings and roads.

2. AUTOMATIC AERIAL TRIANGULATION AND TRANSFORMATION

Traditional Photogrammetry has often 60 percent overlapping between two images. While image sequences taken with digital video camera has the advantages of high overlapping and redundancy of corresponding features, which has a well potential for automatical 3D reconstruction. This will surely improve the speed of 3D reconstruction and also reduce manpower significantly.

Image sequences of 95% overlapping are taken with a digital video camera on a helicopter. The camera is fixed in a stabiliser, which can keep the camera pose stable while the helicopter flying. There are totally three image sequences, one with optical axes looking downward to get the images roofs and roads, other two sequences with oblique view-angles of about 45 degree to get the texture of walls of buildings. The three flying routes are shown in Fig.1.

The most important thing of 3D-model generation is the corresponding problem among images and back-projected spatial data. As we know, ground plans are often not accuracy, so both camera parameters and spatial data must be determined. We adopt AAT to give initial values of camera parameters. Image sequences are processed with VirtuoZo AAT as single strip separately. The output of AAT is a free network taking the first image as coordinate reference, which is different from the world coordinate system. So the camera parameters of the free network need to be transformed into world coordinate system in order to give reliable pose initial values of the image sequences. In this paper, we choose two images near the begin and end part of the image sequence, then some control points are selected from the CAD data and their corresponding image points are given interactively. Thus camera parameters of the two selected images can be obtained with space resection. Then all parameters of the strip can be transformed into world coordinate system. Pass points are also calculated, which can be used in hybrid point-line photogrammetry to ensure the stability of bundle adjustment.

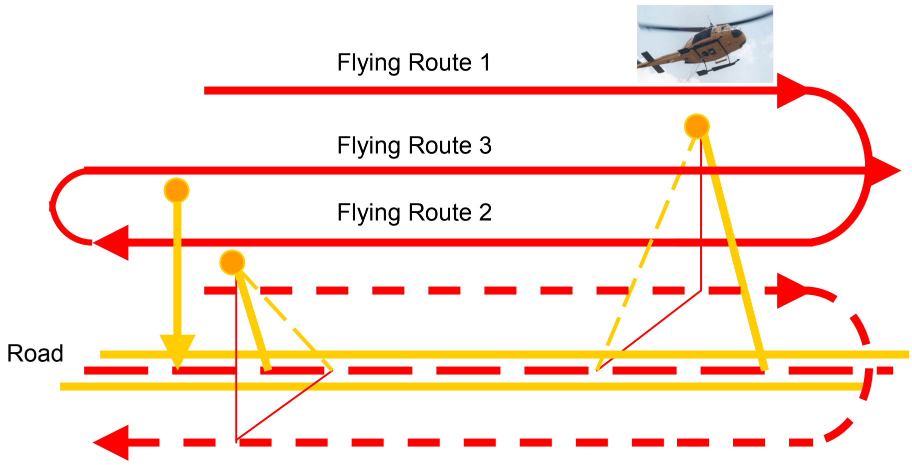

Fig. 1 Three Flying Routes

3. COARSE-TO-FINE BUILDING MODELING

Research on building reconstruction can be subdivided into two large classes, namely automatic systems (Henricsson and Baltsavias, 1997, Baillard et al., 1999) and semiautomatic systems (Grun and Wang, 1998, Brenner, 1999).

Fully automatic systems working on the basis of aerial images are not considered to work reliable enough for practical use, although tremendous progress has been made (Brenner, 2001). Companies actually collect three-dimensional models relying almost exclusively on manual data acquisition, even though the measured models are not very detailed. One of the great disadvantages of systems working on the basis of (Laser-) DSM is the usually low point density of DSM's. This leads to a poor lateral accuracy. Combining DSM's with existing ground plans for automatic reconstruction has shown good results (Brenner, 1999), but DSMs are not always available.

It can be doubted that automatic systems can achieve success rates comparable to human operators within the next few decades (Forstner, 1999), while semiautomatic systems can be used in large practical projects. They can also be taken as kernel of systems where fully automatic algorithms can be incrementally integrated.

Researchers have put forward approaches to model buildings using a small number of primitives—CSG models. But many complex buildings can not be divided into simple CSG models especially in Europe. It is desirable to have semiautomatic systems trying to reduce interactions per building. Moreover, one would like to have a system reconstructing a considerable number of buildings without any operator interference at all.

Nowadays, most city buildings and other man-made objects have 2D ground plans, although they often not as so precise as image resolution. Anyway, they can be used as initial values for automatic reconstruction. So we will use a coarse-to-fine strategy to reconstruct city buildings.

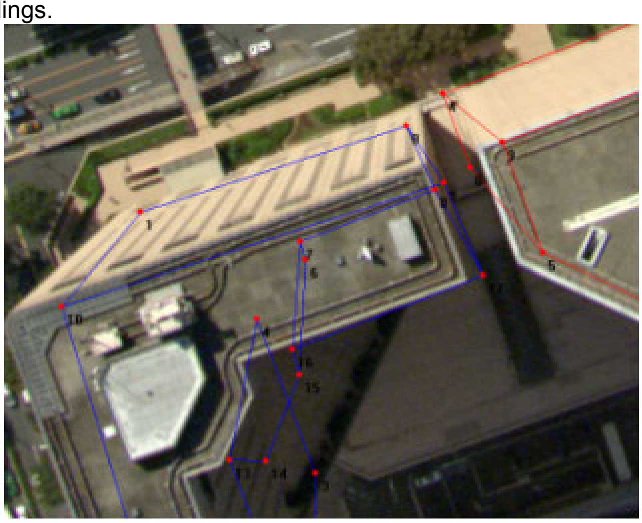

Fig.2. Initial Projections of Buildings

Fig.3. Result Projections with coarse-to-fine strategies

In our system, low accuracy ground plans are back-projected into images according to AAT results to give initial values of buildings, see fig.2. For each building, all related images are automatically obtained from the whole image sequence, and the best image is auto-selected. One image corner of building should be draped to its actual image point coarsely. There also maybe rotation and a scale factor. Building height can also be obtained by draping in image. These operations can be easily done through user interface. After these operations done in the best image, system can automatically match all visible lines of the corresponding building in each image based on least-squares template matching and AAT results.

As we know, not all images have enough building image lines, and the geometry configuration is often very poor. Pass points of AAT are used to ensure stable camera parameters, which can be used to reconstruct other buildings and to generate orthoimages for 3D visualization of terrains. Usually, the model of interested building will significantly improve after bundle adjustment, and errors in ground plans can also be corrected considerably. So there need at most one more user interactive for the interested building. It is shown in fig.3 that projections of other buildings will also more closer to its actual positions in images after another bundle adjustment. Camera parameters of the image sequences are also updated according to bundle adjustment results.

An iterative coarse-to-fine process composed of projection, draping, matching and bundle adjustment is adopted, so user intervene will be decreased very quickly. Finally a large part of buildings can be reconstructed automatically together with precise camera parameters.

4. TEXTURE AND 3D VISUALIZATION

Since one of the most important applications of virtual city models is the generation of realistic visualizations, a proper representation of geometry and texture has to be provided for each building. When all buildings are successfully reconstructed, we can also obtain precise camera parameters of each image. This information can provide a simple back-projection of the reconstructed 3D building to the aerial image, the corresponding image patch for each visible part of a building is available and can be used for texture mapping. The geometry resolution of image patches of walls from downward viewing images is often lower than oblique viewing ones, so the best texture of visible face should be auto-selected and mapped.

Reconstruction of trees in urban areas is also one of the most difficult problems in 3D city model generation. Haala proposed an approach to extract buildings and trees in urban environments based on images and laser altimeter data (Haala 1999), but we know laser-scanning data is not widely available. We generate 3D trees together with orthoimages automatically. Precise camera parameters and image sequences are used to generate orthoimages. Firstly, DEM of the whole region is generated automatically according to matched points in images. Then initial orthoimages and stereo mates are made using camera parameters. DEM can be refined with the orthoimages and stereo mates. Finally, orthoimages including trees are generated based on the refined DEM and camera parameters fully automatically.

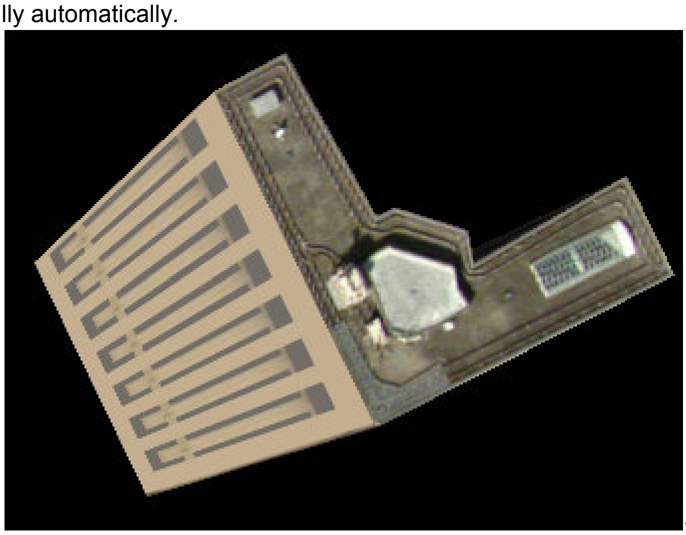

Fig.4. Texture and 3D Visual Model of House

There are often trees in both sides of roads. These trees will occlude roads in downward viewing images and walls in oblique viewing images. For realistic visualization, tree occlusions and cars on the roads must be removed automatically, but transportation marks on roads should be remained. Digital 2D maps of roads are projected into images and corresponding image patches are extracted and processed automatically. Reconstructed buildings are merged into terrains generated from DEM and images. These 3D models are visualized realistically as shown in fig.5.

Fig.5. 3D City Model Visualization

5. REFERENCES

[1] Baillard, C., Zisserman, A. «Automatic Reconstruction of Piecewise Planar Models from Multiple Views». *IEEE Conference on Computer Vision and Pattern Recognition*, 559-565, 1999.

[2] Brenner, C. «Interactive Modeling Tools for 3D Building Reconstruction». In: D. Fritsch & R. Spiller (Eds), *Photogrammetric Week'99*, Wichmann Verlag, Heidelberg, 23-34, 1999.

[3] Brenner, C. «Towards Fully Automatic Generation of City Models». IAPRS Vol. XXXIII, Part B3/1, Comm. III, ISPRS Congress, Amsterdam. 2000, 85-92 2000.

[4] Brenner, C., Haala, N. and Fritsch, D. «Towards fully Automated 3D City Model Generation». Automatic Extraction of Man-Made Objects from Aerial and Space Images III. 2001.

[5] Forstner, W., 1999. «3D-City Models: Automatic and Semiautomatic Acquisition Methods». D. Fritsch & R. Spiller (Eds), *Photogrammetric Week'99*, pp.291-303, Wichmann Verlag, Heidelberg.

[6] Grun, A., Wang, X.. «CC Modeler: A Topology Generator for 3-D City Models». ISPRS Commission IV Symposium on "GIS - Between vision and application", Stuttgart, September 7-10, IAPRS. Vol. 32, Part 4,188-196, 1998.

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

[8] Henricsson, O., Baltsavias, E. «3-D Building Reconstruction with ARUBA: A Qualitative and Quatitative Evaluation». Grun, A., Baltsavias, E., Henricsson, O. (Eds), Automatic Extraction of Man-Made Objects from Aerial and Space Images (II), 65-76, 1997.