Rapidly realizing 3D visualisation for urban street based on multi-source data integration

Zhizhong Kang
Faculty of Aerospace Engineering, Delft University of Technology, Delft, The Netherlands

Zuxun Zhang, Jianqing Zhang
School of Remote Sensing and Information Engineering, Wuhan University, China

Sisi Zlatanova
OTB, GISt, Delft University of Technology, Delft, The Netherlands

ABSTRACT

Streets are important component of cities since they provide the best direct impression of the city. Therefore street scenes are important aspect in 3D modeling. Fast 3D City Modeling from a street level can also be quite important for emergency response by providing realistic, updated, accurate information about accessibility to and from affected areas.
This paper presents a 3D reconstruction approach for rapid 3D visualization from street level, which is based on a combination of vehicle-based image sequence and 2D vector map. The approach consists of two general steps: the rapid reconstruction of facade along the street based on side-shooting vehicle-based image sequence and 2D map, and road texture recovery using. The algorithms presented are verified by experiments on real data set.

Key Words: 3D visualization, street, vehicle-based image sequence, 2D vector map, data integration

1. Introduction

Collecting textures is one of the most critical and time consuming processes in 3D City Modelling (3DCM). Textures for terrain and roof of buildings are often acquired from aerial images and therefore the texture process can be easily automated. The textures for facades, however, are predominantly collected by digital cameras. The image processing such images has usually low efficiency and high labor intensity. Therefore, automation has been always challenging research topic addressed by many.
Many approaches are based on close-range photogrammetry. Pomaska has presented a method to reconstruct 3D model based on multiimage photogrammetry [1]. Recently, spherical and panoramic images are increasingly getting utilized in 3D reconstruction of building. Teller [9] has developed a prototype system for automatic reconstruction using spherical mosaic images, in which camera's position and orientation of each spherical image is first initialized using positioning sensors, then refined through image matching. Haala et al. [7] have presented an automatic method based on mapping of terrestrial panoramic images to the facades of existing 3D building models.
As LIDAR (Light Detecting and Ranging) technology is progressing, 3DCM based on multiple data fusion gains increased attention [2, 3, 6, 8]. However, airborne LIDAR can only capture building roofs but not the facades. This essential disadvantage prohibits their use in photo realistic walk- or drive-through scenes [5].
Another group of researchers focus on video processing. Video's can be quite easily obtained and, consequently, extracted image sequences can be efficiently used to reconstruct textured 3D model. Jun Wu [11] has presented an approach for textured 3D model by fusing helicopter-based video, LIDAR data and 2D vector map. Unfortunately the high cost of helicopter video as well as the dependence on weather conditions and height limit in urban area (resulting in low resolution and large distortion of facade texture), limit the utilization of this approach.
Terrestrial mobile mapping systems have been also used to acquire the complementary ground-level data and to reconstruct building facades. The resolution of textures obtained from vehicle-based sequence is much higher than those obtained from aerial or helicopter image sequence. A recent collection of the mobile mapping literature can be found in Li and Chapman [4] and Tao and Li [10]. A common feature of mobile mapping systems is multi-sensor integration, e.g. CCD, GPS, Inertial Navigation System and dead reckoning. Although some tasks related to calibration and image data processing can be done automatically, a human operator is still needed [5].
The approach presented in this paper, which is also based on a combination of a vehicle-based image sequence and a 2D vector map, provides high automation of the reconstruction process. This paper is subdivided into five parts. Next part presents the steps that complete the reconstruction of the 3D model for the facades. Part three explains the road

texture recovery for the street surface. Experimental results are reported in part four. Part five discusses the results and concludes on further research.

2. 3D reconstruction of facades

The process follows four steps, i.e. image rectification, subsection of image sequence, image mosaic and 3D facade model reconstruction in a semi-automatic way combined 2D vector map.

2.1 Image rectification

The images of facades taken with digital camera are usually very oblique with a large geometric distortion. This requires rectification of all the raw images. The challenge here, usually, is the way of computing orientation parameters (except the coordinates of projective center) using only image information. Since building facades are almost planar, there are two groups of lines in the object space parallel respectively to *X*-axis and *Y*-axis. As a result, only the focus *f* and the three angular orientation parameters φ, ω, k have be calculated applying vanishing point geometry. The facade textures are rectified employing *f* and φ, ω, k with principle point fixed at the center of image. The process of rectification is illustrated as Fig.1. *P* is the original image and P_0 is the rectified image of *P* on the plumb plane.

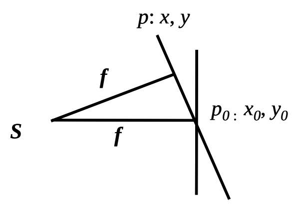


Fig.1 The process of rectification

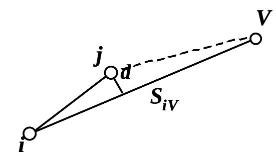


Fig.2 The constrain of Straight lines bundle

The process begin with line extraction on the image. Lines parallel respectively to *X*-axis and *Y*-axis in object space are used to calculate the vanishing points. Each straight line such as line *ij* belonging to a set of parallel lines in object space must pass through the vanishing point as shown in Fig.2. Ideally, the projections of these two group lines in the rectified image should be parallel to *X*-axis and *Y*-axis in object space. Therefore a adjustment approach is used based on constraints of straight lines bundle and known orientation of parallel lines in object space. The interior orientation parameters are fixed in the adjustment.

2.2 Image sequence subsection

To facilitate the reconstruction of the 3D facades, the image sequence is subdivided into segments corresponding to each building on the map. Accordingly, the histogram of projective difference of corresponding points is used for automatic image sequence subsection.

Fig.3 illustrates this process. The origin of the coordinate system is the projective center S_1 of the first image. The plane *XY* is parallel to the facade. The ideal image plane P_0 of the first image (the distance between point S_1 and P_0 is focus *f*) is considered as a public projective plane. As far as the corresponding points on the facade a_1 and a_1' are concerned, the projective difference between their projection points A_1 and A_1' should be very small or even zero. However, projection points A_2 and A_2' of corresponding points a_2 and a_2', which are not on the facade have a large projective difference. Accordingly the 1-D histogram along *x*-axis of the image coordinate system shows the projective difference (Fig.4).

Cross-roads and facades having a large range variance (Fig.5) are represented in the histogram by peak areas. Moreover, the location of range variance in image is indicated by the location of corresponding peak area in the histogram because the histogram is drawn following the moving direction of vehicle. According to the location indicated, images including cross-roads are excluded from image sequence and subsections of image sequence are divided for facades having large variance with each other.

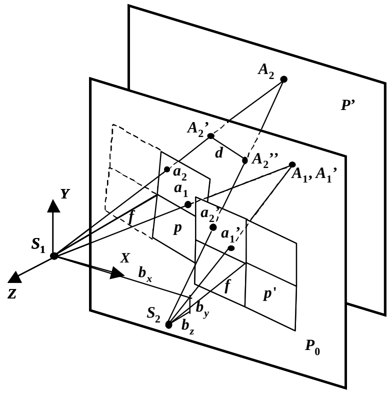
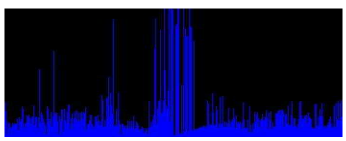
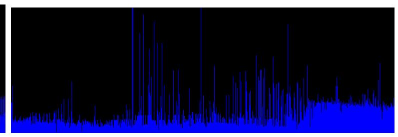
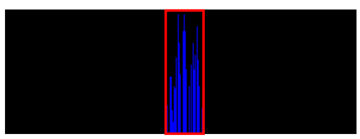


Fig.3 The projective geometry of corresponding points

a Original

a Original



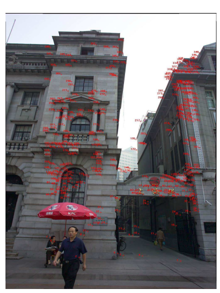

b After filtering

b After filtering

(1) Cross-road

(2) Connected buildings with large range variance

Fig.4 the histogram of projective difference of corresponding points




a. Cross-road

b. Connected buildings with large range variance

Fig.5 Segments with large range variance

2.3 Image mosaic

Usually for large buildings, facade texture is covered by a segment of image sequence as explained the previous section. To be able to complete the entire façade, images have to be combined, which is the next step. Presently, the existing image mosaic software specializes only in processing images with small geometry distortion. However, vehicle-based imagery as discussed in this article has large geometric distortions. This section presents our method for automatic image mosaic after the raw images are rectified onto a vertical plane.

As shown in Fig.3, the facade is considered as a plane and parallel to the public projective plane. Therefore, the relatively spatial relationship of the image sequence is recovered when the corresponding rays of each two images intersect on the public projective plane via adjusting the exterior orientation parameters (EOPs) of each image. Using spatial relationship of the image sequence determined above, a method similar to that of generating ortho-image is employed to generate the whole facade texture. According to the theory of optic projection, the corresponding rays of each image can intersect each other on the public projective plane via adjusting the EOPs of each image. The relative spatial relationship of the image sequence is determined by this least square adjustment.

2.4 Correspondence between facade texture and 2D vector map

After the recovery of facade texture, a 2D vector map is combined to reconstruct the 3D model of building facade. If the facade of a building is approximately planar, only two corresponding corner point pairs need to be selected to link the facade texture with the map. As a result, two parameters scale λ and rotation angle θ between facade texture and map is computed. Roof and height information which is unavailable in the map can be extracted from the facade. Accordingly, the 3D model of facade is reconstructed and geo-referenced to the 2D vector map.

However, the facade of a single building or several connected buildings possibly consists of convex or concave parts which are visible in the vector map. As shown in Fig.6, the corner points (Highlighted by cross) on the convex and concave parts of vector map should be referenced to the facade texture, except for the two corner points highlighted by solid dot.

2.4.1 Corner point extraction

Firstly, two corresponding point pairs A1 and a1, A6 and a6 are selected in vector map and facade texture respectively as shown in Fig.6. The transformation parameters between texture coordinate system and ground coordinate system are computed with these two corresponding point pairs.

a Vector map

b Facade texture

Fig.6 Building with convex or concave parts

The corner points are projected onto facade texture according to λ and θ as point a_i (i = 2,…,5) in Fig.6(b). Afterwards, those points are projected onto raw image sequence. For instance, in Fig.7, projective points in raw images of point a_5 are around the actual corners. However, corner points may not be distinct features in close-ranged images. Furthermore, due to the occlusion between objects with different range, the extraction can also fail. As it can be seen in Fig.7 and 8 that there is always a long vertical line passing the corner of wall and the corner point of interest must lie on this vertical line. The searching for corner point is thereby limited on the vertical line. To extract the vertical line, raw images are firstly rectified onto vertical projective plane and then an image strip, which width is 60 pixels and center is the projective point, is taken out. Consequently, vertical line extraction is implemented in this image strip. As visible in Fig.8, although there might be several vertical lines extracted, the longest line is most likely on the corner of wall. As a result, the longest line highlighted in solid is picked out to search the actual corner point. As image III and IV as be concerned, the vertical lines, extracted from planes with different range, are close to each other so that only the line length is possibly not enough to determine the right line passing corner point.

Fig.7 The projective points of corner point in raw image sequence

To tackle this problem, a strategy to eliminate false accepted line is implemented based on the relative spatial relationship of image sequence acquired after image mosaic. As shown in Fig.8, vertical lines are firstly extracted from the raw image (e.g. I) in which the geometric difference between planes with variant range values is most significant since it is most likely to pick up the correct vertical line from this image. In this image, the longest vertical line is selected and projected onto other adjacent images according to the relative spatial relationship of image sequence. The projective line in image IV is highlighted in square dot and can be used as predicted position of the correct vertical line (Fig 9). As a result, an image strip, which width is narrowed to 40 pixels and center point is the center of projective line, is taken out for vertical line extraction. Although the predicted position is close to the actual vertical line, as Fig.8, the distance is only 6 pixels between the dash line and actual solid line in the while rectangle and moreover the dash line is long as well. Under this circumstance, the dash line is possibly to be falsely accepted. Therefore, the selected line is projected into image I. As Fig.10, the distance is still small between the correct projective line from image IV (highlighted in dash) and the actual vertical line in image I (highlighted in solid). The distance, however, becomes large between the false dash line and the actual solid line. Obviously, the falsely accepted line can be eliminated by this strategy.

After the extraction of vertical lines passing corner points, the line between corner points a_1、 a_6 in Fig.6 is projected into raw images and the actual corner points are the intersected points between extracted vertical lines and projective line of a_1a_6 in corresponding raw image. All extracted and selected corner points are used to calculate parameters scale λ and rotation angle θ .

Fig.8 Vertical lines extracted

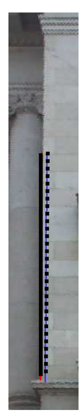


Fig.9 Line prediction

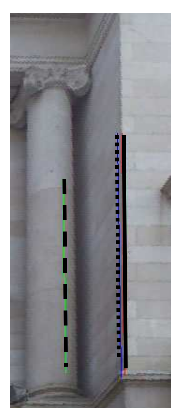


Fig.10 Error line elimination

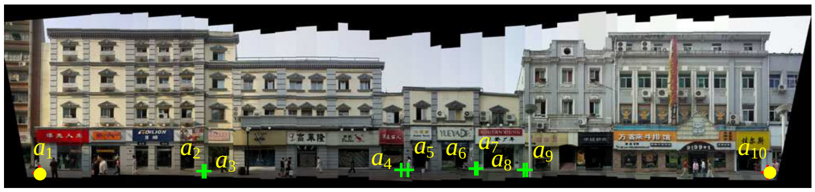
2.4.2 Iterative process

Since the scale factor λ and rotation angle θ are firstly computed with the two corresponding point pairs selected, the further the cross points on the convex and concave parts on the texture away from the solid dot corner points, the larger deviation is between projective and actual points in facade texture, e.g. projective point a_4~a_7 in Fig.11 (a). The iterative strategy is thereupon employed to tackle this problem.

Based on the method presented above, points a_2, a_9 closest to solid dot corner points are firstly extracted from raw images and then λ and θ are recalculated combined with extracted points and selected corner points. According to the

transformation parameters recalculated, point a_3 and a_8 are extracted. This iterative process repeats till all corner points are extracted. In Fig.11 (b), large deviation between projective point a_4~a_7 and actual corner points in facade texture is eliminated after iterative process.
Using all of the extracted and selected corner points, combined with facade texture, the 3D model of facade is reconstructed and geo-referenced to the 2D vector map.

a Before iterative process



b. After iterative process

Fig.11 Projective points

Fig.12 Image sequence along the optic axis

3. Road texture recovery

The image sequence illustrated in Fig.12 is taken by digital camera moving along the optical axis. Since the whole road texture instead of facade texture can be covered, this kind of image sequence is a good source for road texture recovery. In this section a method is therefore presented to recover road texture using this data source. It consists of two steps: image rectification to level plane and image mosaic.

3.1 Image rectification to level plane

Similar to the facade texture recovery, raw images need to be rectified onto level plane to cover the texture close to the real road surface.
As shown in Fig.13, plane $\boldsymbol{p}$ denotes the raw image, $\boldsymbol{p_0}$ denotes the image rectified to vertical plane, $\boldsymbol{P}$ denotes the image rectified to level plane. As Fig.12, because the raw image is taken along the optical axis, the ideal image plane of raw image is the vertical plane. Therefore, the process of rectification is $\boldsymbol{p} \rightarrow \boldsymbol{p_0} \rightarrow \boldsymbol{P}$, namely the raw image is firstly rectified onto the vertical plane and then rectified onto level plane.
A least square adjustment system based on the constraint of straight lines bundle is used to compute focus f and three angular orientation parameters φ, ω, k . Then raw images are rectified onto level plane using those computed parameters. As we know, the larger is the depth of road surface, the larger geometric distortion has the rectified image. As a result, only the region with the size of 3000 pixels$\times$2000 pixels which has smallest geometric distortion in the rectified image is used to recover road texture (Fig.14). Image mosaic is necessary to acquire the whole road texture of street with high resolution.

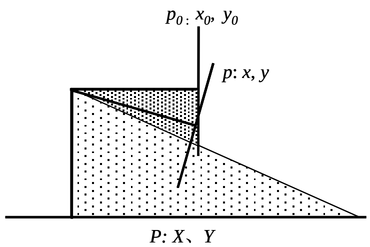


Fig.13 Rectification process

Fig.14 Rectified road texture

3.2 Mosaic for road texture

The road surface can be considered as a plane and parallel to the public projective plane. The projective geometry of corresponding points of the image sequence along the optical axis is illustrated in Fig.15. The level image plane P_0 of the first image is considered as public projective plane. According to the theory of optic projection, the corresponding rays S_1a_1 and S_2a_2 can intersect at point A_0 on plane P_0 via adjusting the EOPs of each image. A least square adjustment is employed to calculate the EOPs by which all of coordinate difference between correspondingly projective points is least square. Using the EOPs computed, a method similar to that of generating ortho-image is employed to generate the whole road texture of street (Fig. 16).

As presented in Section 2, the 3D model of street facade was reconstructed and geo-referenced to the 2D vector map, therefore integrated with the road texture of street recovered the 3D visualization of walk-through street scene is realized.

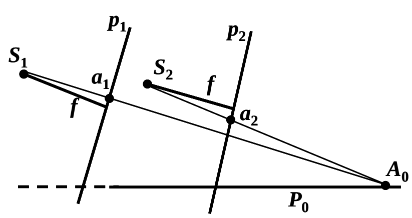


Fig.15. The projective geometry of corresponding points (along the optical axis)

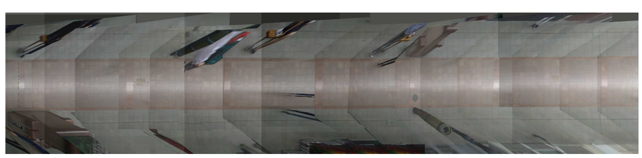


Fig.16 mosaic road texture

4. Experimental results

Experiments were implemented on two image sequences, i.e. side-shooting (Fig.17) and along the optic axis (Fig.12), taken by digital camera KODAK PROFESSIONAL DCS Pro SLR/n along the street. For side-shooting, the distance between photographic center and building is about 10m. The image size is 3000 pixels×4500 pixels. To fix the interior parameter during the process of taking photos, the focal length is set to infinitude and the function of automatic focalization is turned off. Actually, it's not necessary to process the imagery with such large image size, therefore, the raw images were compressed and image size is reduced to 1500 pixels×1000 pixels.

Fig.17 Side-shooting image sequence

Automatic image sequence subsection was implemented using the method presented in Section 2.2. After the subsection of image sequence, every building has a corresponding image segment to recover facade texture. As presented in Section 2.3, image mosaic is required to generate the whole facade texture. For comparison, facade texture is mosaic from close-ranged digital camera images respectively by the existing image mosaic software and the strip method presented in Section 2.3. The results of existing image mosaic software are shown respectively in Fig.18. Obviously, the existing software tools of image mosaic are not applicable for processing images with large obliquity of the posture of digital camera.

a. ImageSuite *b*. PhotoStitch

Fig.18 Mosaic obtained using existing software

The average time consuming for the rectification of a single image is 30 seconds. As the building facade with 82m length in Figure as be concerned, the corresponding image segment consists of 36 images. Total time of rectification is 18 minutes. For image mosaic, there are 216 unknown parameters in the adjustment system of strip method and the time consuming is 10 seconds. The whole process is fully automatic and fast. Moreover, as Fig. 19, the facade texture generated from 36 images has a very high resolution and close to the seamless mosaic.

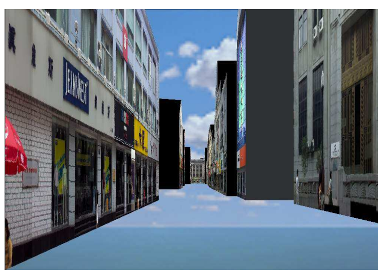
Fig.19 Whole facade texture

Combined the facade texture and 2D vector map, the 3D model of the facade was reconstruct as shown in Fig.20 (*a*). For the facade consists of convex or concave parts, 3D model was reconstructed as Fig.20 (*b*) using the semi-automatic algorithm presented in Section 2.4.

a. Planar facade

b. Buildings with convex or concave parts

Fig.20 3D facade models



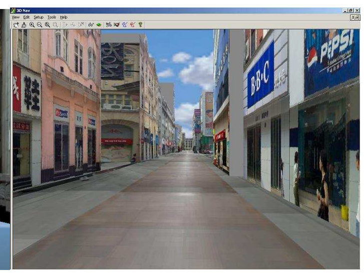


Fig.21 3D model of street   Fig.22 3D walk-through scene of street

Finally 3D models of 28 building facades along the pedestrian street were reconstructed and geo-referenced to 2D vector map (Fig.21). Using the method presented in Section 3, the whole road texture of street was recovered from 255 images along the optic axis. Integrated the road texture recovered with 3D models of building facades, the 3D model of walk-through street scene is reconstructed (Fig.22). Since terrestrial close-ranged images have high resolution and moreover are taken along the street, the walk-through street scene is close to the real impression of people on the street.

5. Conclusion

In this paper, an approach was presented that allows for fast reconstruction and 3D visualization for walk-through street scene. The image sequences (side and along the optic axis shooting) are collected by camera mounted on a vehicle. The presented approach has numerous advantages compare to other approaches:
Because the acquisition of image sequence is vehicle-based and moreover the recovery of building facade texture which is always low efficiency to existing methods realizes automation in our method, the whole process from data acquisition to 3D model reconstruction is implemented in a rapid and high efficient way.
Image sequence along the optic axis is an ideal source to recover road texture with high resolution. The texture recovery can be completed by image rectification and mosaic in an efficient way. Terrestrial close-ranged images have high resolution and since they are taken along the street, it is possible to create a walk-through 3D scene, which is very close to the reality. The approach presented in this paper is applicable also for streets with high density of buildings.
This approach is quite appropriate fast reconstruction and 3D visualization is emergency response (e.g. flood or earthquake). The process is fully automatic and very fast. Since the texture is acquired from images, the scene should have good visibility (with some exceptions, e.g. smoke). The building facades should be roughly or partly complete so that the facade model can be reconstructed from image sequence and 2D vector map. The image sequence can be acquired

by a person instead of a vehicle moving along the street if the road is blocked or damaged. The developed prototype system can be further optimized to allow for near real-time reconstruction and visualization. First estimations of building damages can be then send to rescue teams via internet.
Future work will concentrate on 3D model reconstruction for more complex facade and improvement of availability for disaster management.

Acknowledgements

This research was supported by the National High Technology Research and Development Program of China (863 Program) with the serial number 2006AA12Z220.

References

1. Grunter Pomaska (1996) Implementation of digital 3Dmodels in building surveys based on multiimage photogrammetry[A]. In: International Archives of photogrammetry and Remote sensing[C],Vienna, Vol.XXXI, Part 5,pp 487-492.
2. Impyeong Lee, Yunsoo Choi (2004) Fusion of Terrestrial Laser Scanner Data and Images for Building Reconstruction [A], Proc. of ISPRS [C], Vol. XXXV, part B5.
3. KAZUO ODA, TADASHI TAKANO, TAKESHI DOIHARA, RYOSUKE SHIBASAKI (2004) Automatic building extraction and 3-D city modeling from LIDAR data based on hough transformation [A]. Proceedings of XXth ISPRS Congress [C],Vol. XXXV, part B3.
4. Li, J. and M. Chapman (eds.) (2005) Mobile Mapping Systems, Special Issue of Photogrammetric Engineering & Remote Sensing, Vol. 71, No. 4.
5. Li, J. and M. Chapman (2007) Terrestrial mobile mapping towards real-time geospatial data collection, to be included in (eds.) Zlatanova and Li, GI technology for emergency response, ISPRS book series (to be published 2007)
6. Maas, H.-G., and G. Vosselman (1999) Two algorithms for extracting building models from raw laser altimetry data, ISPRS Journal of Photogrammetry & Remote Sensing, 54(2–3):153– 163.
7. Norbert Haala, Martin Kada (2005) Panoramic Scenes For Texture Mapping Of 3d City Models [A], Proceedings of the ISPRS working group V/5 'Panoramic Photogrammetry Workshop' [C], IAPRS Vol. XXXVI-5/W8, on CD.
8. Schwalbe, E., H.-G. Maas, and F. Seidel (2005) 3D building model generation from airborne laser scanner data using 2D GIS data and orthogonal point cloud projections, Proceedings of ISPRS WG III/3, III/4, V/3 Workshop on Laser Scanning 2005, Enschede, the Netherlands, September 12-14, pp 209-214.
9. Seth Teller (1998) Toward Urban Model Acquisition from Geo-Located Images[A], In Proceedings of Pacific Graphics , pp 45--51.
10. Tao, C. V. and J. Li (eds.) (2006) Advances in Mobile Mapping Technology, ISPRS Book Series, Taylor & Francis.
11. Wu Jun (2005) Research on Rapidly Reconstructing Texture for Facades in 3D City Modeling, ACTA GEODAETICA ET CARTOGRAPHICA SINICA, 34(4): 317-323.