

A Framework for Building Reconstruction Based on Data Fusion of Terrestrial Sensory Data

Impyeong LEE* and Yunsoo CHOI**

Abstract

Building reconstruction attempts to generate geometric and radiometric models of existing buildings usually from sensory data, which have been traditionally aerial or satellite images, more recently airborne LIDAR data, or the combination of these data. Extensive studies on building reconstruction from these data have developed some competitive algorithms with reasonable performance and some degree of automation. Nevertheless, the level of details and completeness of the reconstructed building models often cannot reach the high standards that is now or will be required by various applications in future. Hence, the use of terrestrial sensory data that can provide higher resolution and more complete coverage has been intensively emphasized. We developed a fusion framework for building reconstruction from terrestrial sensory data, that is, points from a laser scanner, images from digital camera, and absolute coordinates from a total station. The proposed approach was then applied to reconstructing a building model from real data sets acquired from a large complex existing building. Based on the experimental results, we assured that the proposed approach can achieve high resolution and accuracy in building reconstruction. The proposed approach can effectively contribute in developing an operational system producing large urban models for 3D GIS with reasonable resources.

Keywords : Data Fusion, Building Reconstruction, Registration, Modelling

1. Introduction

The next generation applications such as Telematics, LBS (Location Based System), and Ubiquitous has been rapidly developed in recent years. As the most crucial basis for such applications, the importance of 3D GIS has been increasingly emphasized. One of the important components for 3D GIS is considered as urban models; and hence their acquisition and periodical update are inevitable to maintain operational 3D GIS. For reference, the current status and perspectives on 3D GIS are presented by Stoter and Zlatanova (2003) and Zlatanova et. al. (2002).

Urban models mainly contain models of artificial structures such as buildings and roads. Many studies have then focused on extracting buildings and roads from sensory data in automatic or semi-automatic ways. Particular interests have been in building reconstruction, which attempts to generate building models

usually from sensory data. The building models usually include the geometric descriptions (shapes and positions) and radiometric descriptions (texture) of a building. The sensory data used for building reconstruction have been mostly aerial or satellite images, more recently airborne LIDAR data, or the combination of these data (Halla and Brenner, 1999; Maas and Vosselman, 1999; Jung, 2004).

The previous studies on building reconstruction have developed some competitive algorithms with reasonable performance and some degree of automation. Nevertheless, the level of details and completeness of the buildings reconstructed using such algorithms often cannot reach the high standards that is now or will be required by 3D GIS in future. This problem mainly originates from the limitation of the sensors themselves. That is, since most sensors acquire the data in far-range, for example, from air or space, they cannot provide the data of sufficiently high

*Dept. of Geoinformatics, University of Seoul, Seoul 130-743, Korea (E-mail : iplee@uos.ac.kr)

**Dept. of Geoinformatics, University of Seoul, Seoul 130-743, Korea (E-mail : choiys@uos.ac.kr)

precision and complete coverage. Hence, the use of close-range sensors has been recommended in such cases.

The most representative close-range sensors are digital cameras and laser scanners nowadays. The cameras produce images while the scanner points. The images provide better information about the edges and corners of buildings while the points about the facets. In addition, the reconstruction processes starting from images were not easily automated while those from points more possibly automated. More elaborate comparison results are presented by Ackerman (1999), Baltsavias (1999), and Schenk (1999). Hence, to benefit from such complementary properties, it is more desirable to employ a fusion approach that use both points and images simultaneously for building reconstruction.

Based on the careful review on the research trends and the needs of applications available now or to be flourished in future, we argue that operational systems should be developed to perform building reconstruction from images and points acquired by close range sensors in interactive or nearly-automatic ways. Hence, the purpose of this research was to perform a pilot study toward the development of such systems. We thus established a conceptual framework for such systems and the core processes constituting the framework. We then validated this proposed framework and processes by applying them to reconstructing a building model from real data sets acquired from a large complex existing building.

This paper introduces the proposed framework consisting of two main stages, data acquisition and data

processing in section 2. Section 3 presents the experimental results from the application to real data sets. Finally, this paper concludes with summary and some remarks in section 4.

2. The Building Reconstruction Framework

We propose a framework to generate precision building models from terrestrial sensory data. As shown in Fig. 1, this framework includes two main stages, *data acquisition* and *data processing*. Data acquisition involves collecting three types of data about a target (a building). These data are surface points, images, and reference points acquired by a laser scanner, a digital camera, and a total station, respectively. Data processing involves deriving a precision building model from these sensory data. The model retains geometric information (shapes and positions) in absolute coordinates with radiometric information (texture).

2.1 Data Acquisition

2.1.1 Sensor overview

Using a laser scanner, we can acquire numerous points densely sampled from the surfaces of a building. The laser scanner transmits laser pulses to the target and receives the return pulses reflected at a surface of the target. It then computes the travelling time of a laser pulse and converts it into the range between the scanner and the target. The range is combined with the direction of the pulse recorded by the scanner itself to generate three dimensional coordinates of the point on the target at which the pulse is reflected. The coordinates are

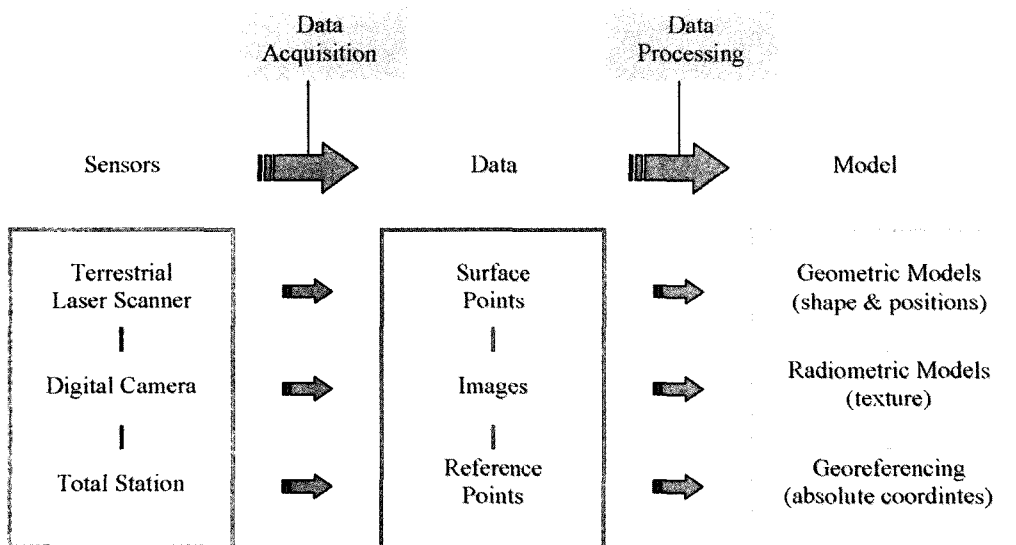


Fig. 1. Framework for generating precision building models from terrestrial sensory data.

expressed in a coordinate system fixed to the laser scanner. The coordinate system depends on the orientation and position of the scanner and thus is not an absolute coordinate system fixed to the Earth.

Using a digital camera, we can acquire color images of a building. This camera is not necessarily a metric camera retaining high resolution and accuracy. A general purpose digital camera must be sufficient since the images are used for obtaining only texture information. To derive precise geometric information from the image, we should first perform camera calibration processes. A promising algorithm for the calibration is presented by Habib et. al. (2002).

Using a total station, we can acquire the accurate coordinates of reference points in an absolute coordinate system. A prism rod is erected at each reference point to facilitate a convenient measurement of its location using the total station. The measured reference points should be also identifiable from the point sets and images acquired by the laser scanner and digital camera so that they can be used for geo-referencing these data.

2.1.2 Sensor Configuration

Sensor configuration considers positioning sensors and locating reference points. A field of view of a sensor is always limited and thus the data set acquired at a position can cover only partial area of a building. Hence, the sensors should be carefully positioned so that the entire facets of the building can be covered. Data sets obtained at adjacent positions should include sufficient overlaps for their registration. Reference points are located so that each data set can include four reference points that do not lie on a straight line. Also, at least three reference points should be precisely measured by a total station to provide absolute coordinates. Two points among them are overlapped by an

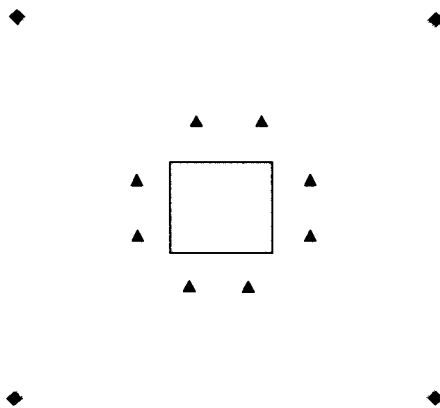


Fig. 2. An example of sensor configuration.

adjacent data set. Fig. 2 shows an example of sensor configuration, where the diamonds indicate the locations of the laser scanner and digital camera and the triangles those of the reference points.

2.2 Data Processing

Data processing for generating a precision building model from the acquired sensory data involves two main processes, registration and modelling.

2.2.1 Registration

Each data set is defined in its own local coordinate system. To combine such data sets, we should determine the relationships between the coordinate systems. Registration aims at defining about an absolute coordinate system the local coordinate system in which each data set (either laser scanner point sets or digital camera images) is expressed. Registration is based on the reference points.

We establish a model graph for the point sets acquired at all the positions, as shown in Fig. 3. The model graph includes each point set as a node and represents the existence of overlap between two point sets as an arc. Each arc incorporate a transformation represented as T_{ij} , which establishes relationship between two coordinate systems in which two point set are defined. It makes it possible to convert a point set into the coordinate system of another points and vice versa. The final goal of the registration is to convert each point set into an absolute coordinate system denoted as G. This transformation is denoted as T_{iG} . Both kinds of transformation, T_{ij} and T_{iG} are classified into 3D similarity (or called rigid-body) transformation that have three parameters for rotation, (ω, ϕ, κ) and

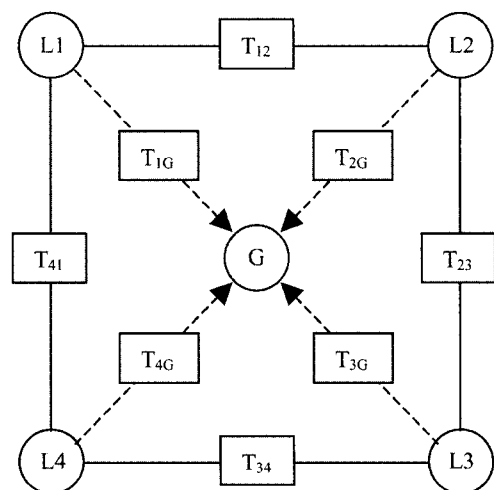


Fig. 3. An example of a model graph for point sets.

three parameters for translation, (x_t, y_t, z_t) , defined as

$$\begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} = R(\omega, \phi, \kappa) \begin{bmatrix} x_G \\ y_G \\ z_G \end{bmatrix} + \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix}, \quad (1)$$

where $R(\omega, \phi, \kappa)$ is a 3D rotational matrix defined with three rotational angles, ω , ϕ , and κ with respect to x, y, and z-axis, respectively.

We establish a model graph for the images acquired at all the positions, as shown in Fig. 4. The model graph includes each image as a node and represents the existence of overlap between two images as an arc. Each arc incorporate a transformation represented as R_{ij} , which establishes relative orientation between two images (Ii, Ij). The final goal of the registration is to establish the exterior orientation E_{iG} between each image and an absolute coordinate system. With the exterior orientation, we determine where a point in the object space appear in an image based on the collinearity equations

$$\begin{aligned} x_i &= x_p - c \cdot \frac{(X_G - X_o)r_{11} + (Y_G - Y_o)r_{12} + (Z_G - Z_o)r_{13}}{(X_G - X_o)r_{31} + (Y_G - Y_o)r_{32} + (Z_G - Z_o)r_{33}} \\ y_i &= y_p - c \cdot \frac{(X_G - X_o)r_{21} + (Y_G - Y_o)r_{22} + (Z_G - Z_o)r_{23}}{(X_G - X_o)r_{31} + (Y_G - Y_o)r_{32} + (Z_G - Z_o)r_{33}} \end{aligned} \quad (2)$$

where (x_i, y_i) is the image point corresponding to the ground point (x_G, y_G, z_G) defined in an absolute coordinate system; (X_o, Y_o, Z_o) are the coordinates of the position of the camera; and (r_{ij}) is an element of the rotational matrix composed by three rotation angles (ω, ϕ, κ) indicating the orientation of the camera.

We estimate the transformation parameters associated

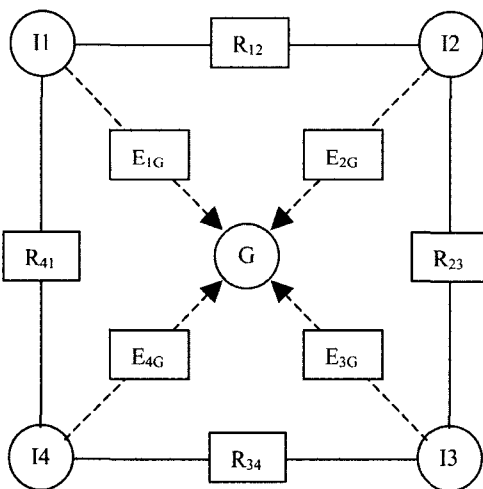


Fig. 4. A model graph for images.

with the registration of point sets and images using block adjustment. The main characteristic of the block adjustment is simultaneous determination of the transformation parameters of the entire data sets rather than sequential determination of the parameters of each individual data set. We establish a set of non-linear equations associated with the entire data sets. This set includes the transformation parameters as unknowns and the coordinates of some points in the data sets as observations. Some points among the points used as observations are also measured by a total station to provide their coordinates in an absolute ground coordinates. These points serve as control points that relates the data expressed in local coordinate systems to the ground coordinate system. In case of the example in Fig. 2, four sets of 3D similarity transformation parameters for the point sets and four sets of exterior orientation for the images are to be estimated.

2.2.2 Building Modelling

Modelling aims at extracting a three dimensional precision model of a building from the geometrically registered point sets and images. The overall modelling processes are presented in Fig. 5. We first extract primitive features from the raw data and produce derived features from the primitive features.

From points, we can extract patches by grouping points into sets of spatially coherent points, where a patch is a cluster of points that can be smoothly approximate to a plane and the overall outline of the cluster can be used as the boundary of the patch. Then, we establish the adjacency among the patches and derive edges by intersecting each pair of adjacent patches.

From images, we can extract edges by applying an edge operator (ex. Canny operator) to derive edge pixels, linking these pixels to generate edge lines, and selecting long straight lines among them. Then, we derive patches by grouping the edges to construct closed polygons.

It is noticeable that the edges and patches can be extracted independently from points and images. So, we can refine the edges and patches by comparing those generated from points with those from images. The refined edges and patches are grouped into a polyhedron representing a geometric (shape) model. Each facet of the polyhedron is draped with the texture derived from images. This draped polyhedron is the final result of these building modelling processes.

Particularly, extracting patches from points and grouping the patches into polyhedron models are based

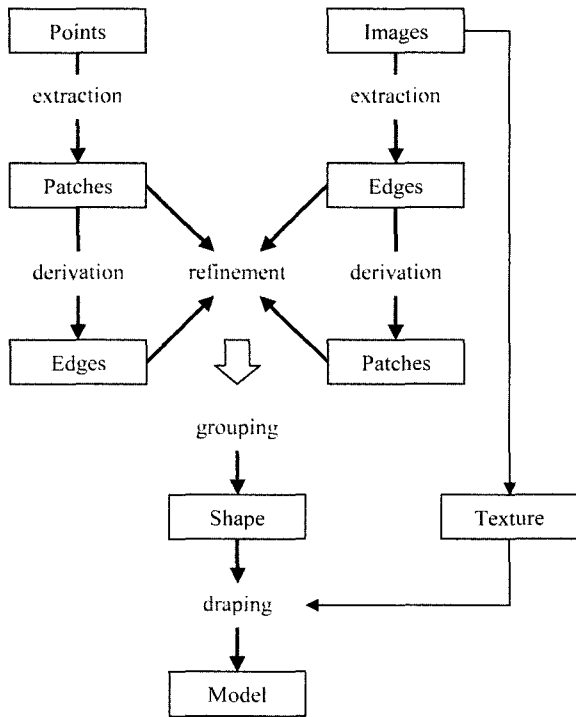


Fig. 5. The overall processes of building modelling.

on the three dimensional perceptual organization processes developed by the author in the previous research (Lee and Schenk, 2001a; Lee and Schenk, 2001b; Lee and Schenk, 2002; Lee, 2002).

3. Experimental Results

The proposed framework and approach was not fully implemented yet and hence only limited parts were verified based on the relevant experiments. We summarize the results from acquisition, processing (mainly focusing on registration) and evaluation.

3.1 Test Site and Acquisition Results

Fig. 6 shows the building used for the experimental results. This building comprises a complex structure retaining different appearances for each side.

We used a terrestrial laser scanner (mode number: ILRIS-3D) manufactured by Optech. Based on its specifications, about ± 10 mm of accuracy for the acquired point coordinates were expected. Table 1 summarizes the results from the acquisition by the laser scanner. We scanned the building at 44 different positions, covering the entire building facets with a nominal point spacing of about 6 cm and providing more than six million points.

We used a digital camera (mode number: F828)

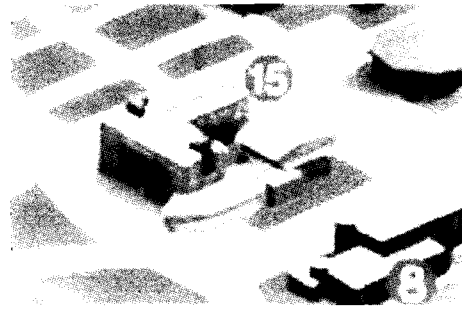


Fig. 6. An air view of the building (denoted with 15) used for the experimental tests.

Table 1. Results of acquisition by the laser scanner.

Number of point sets	44
Average point spacing	5.96 cm
Average range	38.39 m
Number of points	6,454,580
File Size	100.5 MB

Table 2. Results of acquisition by the digital camera.

Number of images	29
Dimension	2592 × 1944
File Size	16.3 MB

manufactured by Sony. Table 2 summarizes the results from the acquisition by the camera.

3.2 Processing and Evaluation Results

The 44 points sets and 29 images acquired at different positions were integrated using the registration processes. The residuals resulted from the registration are within the acceptable range, ± 10 mm in RMS (root-mean-squares). These registered data were used to generate a model using a semi-automatic means implemented based on the proposed framework.

The derived polyhedron model was evaluated by comparing the model with the original planning model of the building. A floor plan was extracted from each model, as shown in Fig. 7. We compared the length of every corresponding straight segment. The histogram of the differences in lengths resulted from this comparison is shown in Fig. 8. The standard deviation of the differences is about ± 37 cm. This magnitude is much larger than the value (± 1 cm) we expect based on the specifications of the laser scanner. This indicates that the differences originate from other sources than the random errors associated with the laser scanner measurements. Also, this implies that there are significant differences between the planning model and

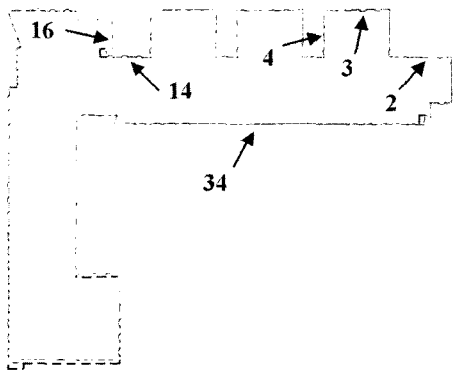


Fig. 7. Comparison of the floor plan extracted from the derived model (red solid polygon) with that from the planning model (blue dashed polygon).

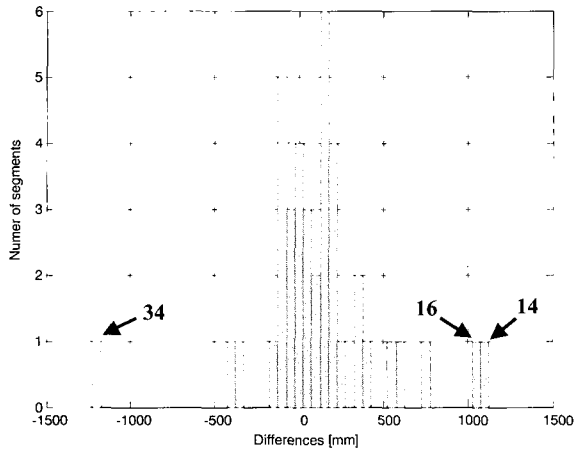


Fig. 8. Histogram of differences in the lengths of segments.

Table 3. Comparison of lengths of three segments. (unit : mm)

No.	Planning Model	Derived Model	Total Station
2	9,975	10,013	10,030
3	7,500	7,344	7,390
4	10,350	10,404	10,400

derived model. To clarify which one is correct, we directly measured three segments (2, 3, 4) using a total station. These results are described in Table 3. Every difference between that from the derived model and that from the total station is less than 5 cm. The values from the derived model are more consistent with the total station results than those from the planning model. This is reasonable observations since the planning model is not usually perfectly kept during the construction of a building. Furthermore, three segments (14, 16, 34)

noticeably show the difference of larger than 1 m, as indicated in Fig. 8. This must be also due to the construction errors.

4. Conclusions

We proposed a conceptual framework and the core processes for building reconstruction from images and points acquired by laser scanners, digital cameras, and total stations. The core processes were then partially implemented and applied to real data sets to evaluate the proposed framework and processes. The evaluation results support the promising performance of the proposed processes in terms of the precision and accuracy of the derived building models. The limitation of this study is a partial implementation and experiment of the proposed entire framework and hence we will establish a practical semi-automatic system based on this framework in near future.

Acknowledgements

This research was based on “2003 high-tech equipment supports” from The University of Seoul. We would like to thank Seokjae Sa, Euijong Oh, Changho Yu, Donghyun Kim and other members of the geodesy and photogrammetry group for acquiring and processing the data.

References

1. Ackermann, F. (1999). Airborne laser scanning – present status and future expectations. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54 (2-3), pp. 64-67.
2. Baltsavias, E. (1999). A comparison between photogrammetry and laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54 (2-3), pp. 83-94.
3. Habib, A., Shin, S., and Morgan, M. (2002). New Approach for calibrating off-the-shelf digital cameras. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Graz, Austria, Vol. XXXIV, Part 3A, pp. 144-149.
4. Halla, N. and Brenner, C. (1999). Extraction of buildings and trees in urban environments. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54 (2-3), pp. 130-137.
5. Jung, F. (2004). Detecting building changes from multi-temporal aerial stereopairs. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58 (3-4), pp. 187-201.
6. Lee, I. (2002). *Perceptual Organization of Surfaces*. Ph.D. Dissertation, Dept. of Civil and Environmental Engineering and Geodetic Science, The Ohio State University, Columbus, OH.
7. Lee, I. and Schenk, T. (2001a). Autonomous extraction

- of planar surfaces from airborne laser scanning data. In: *ASPRS Annual Conference*, St. Louis-MO.
8. Lee, I. and Schenk, T. (2001b). 3D perceptual organization of laser altimetry data. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Annapolis-MD, Vol. XXXIV, Part.3/W4, pp. 57-65.
 9. Lee, I. and Schenk, T. (2002). Perceptual organization of 3D surface points. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXXIV, Part 3A, pp. 193-198.
 10. Maas, H. -G. and Vosselman, G. (1999). Two algorithms for extracting building models from raw laser altimetry data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54 (2-3), pp. 153-163.
 11. Schek, T. (1999). Photogrammetry and laser altimetry, In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, La Jolla, CA, Vol. XXXII, Part.3-W14, pp. 3-12.
 12. Stoter J. and Zlatanova, S. (2003). 3D GIS, where are we standing, In: *Joint Workshop on Spatial, Temporal and Multi-Dimensional Data Modelling and Analysis*, Quebec city, Canada, 6p.
 13. Zlatanova, S., Rahman A. A. and Pilouk, M. (2002). 3D GIS: current status and perspectives, In: *the Joint Conference on Geo-spatial theory, Processing and Applications*, Ottawa, Canada, 6p.