AUTOMATIC RECONSTRUCTION OF LARGE-SCALE VIRTUAL ENVIRONMENT FOR INTELLIGENT TRANSPORTATION SYSTEMS SIMULATION

Khairil Azmi, Shintaro Ono, Masataka Kagesawa, Katsushi Ikeuchi
Institute of Industrial Science, the University of Tokyo
Ee405, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505 JAPAN
Tel: +81 3 5452 6242 Fax: +81 3 5452 6244
Email: {kazmi} | {onoshin} | {kagesawa} | {ki}@cvl.iis.u-tokyo.ac.jp

ABSTRACT

Creating a large-scale virtual environment, such as that of an entire town or city, is one of the most important applications in ITS (Intelligent Transportation Systems). This paper introduces an automatic approach for the reconstruction of large-scale 3D virtual environment intended for use in ITS related simulations, in which a 3D virtual city model generation system and a time series traffic and air pollution simulation combined, whereas a simple and relatively robust approach based on shadow analysis for 3D object automatic reconstruction is proposed.

INTRODUCTION

City model is the most well known example of virtual environment. Interest in 3D city model has risen significantly in the past years. Being one of the most important themes in research and development on Intelligent Transportation Systems (ITS), traffic simulations have been thought of being one of its major application areas.

The development of tools for the reconstruction of 3D city models has been a topic of intense research for the past years. Modeling from images is a classic problem in computer vision and remote sensing. Modeling from aerial images have several advantages over modeling from terrestrial images since aerial images provide accurate building footprints and roof heights whom terrestrial ones do not provide, and they also can be rectified into orthoprojections, which facilitate the merger of multiple images to cover large geographic areas. For this reason, aerial images are adopted as the initial input in our method of obtaining 3D information for reconstructing large-scale 3D city models.

The complexity and quantity of data needed to represent large-scale city models makes fully automatic techniques highly desirable. Still, many barriers to fully automatic systems exist. For example, image noise, lighting conditions, occlusions, and scene complexity complicate segmentation—the identification of buildings in an image. Moreover, 3D reconstruction of complete buildings is difficult when a single building consists of complex substructures. However, when physical rough factors of the structures (e.g. height of buildings, range between buildings, etc) are thought of being more essential than reality and aesthetic aspects (for example, in the case of a 3D city model reconstructed for the use in air pollution simulation), there is no urgent need to obtain and process the data of textures and other complex details of the structures. Hence, a fully automatic reconstruction system then can be brought into reality.

On the other hand, air pollutants quantity released by vehicles is parallel to traffic quantity. It is clear that congested roads are more polluted than empty ones for nearly every single motor vehicle releases pollutants into the air, while pollution intensity itself is not constant since traffic quantity simultaneously changes time to time. And there is no doubt that one of the most significant factors that determine how the released pollutants should disperse in the air is the height of structures surrounding the roads.

This paper proposes a method for automatically reconstructing 3D city model for the use in traffic simulation and visualization of time series change of air pollutants dispersion, in which height of structures is the key information to acquire. The next section describes a comparison to related work, which is then followed by an explanation on 3D city model automatic reconstruction method based on shadow analysis of aerial imagery. This section is then followed by a short description of the proposed 4-dimensional traffic and air pollution simulation as an application area of the obtained city model.

RELATED WORK

Extensive research has examined the use of single and multiple aerial images in city modeling [1]. Lin et al. use monocular aerial images from a general viewpoint to detect urban buildings and construct 3D shape descriptions of them [2]. Vanden et al. use geometric and projective constraints to extract 3D models from a single image [3]. Baillard et al. use multiple aerial images and stereo algorithms to extract 3D models [4].

On the other hand, however, parallel to the rapid development of remote sensing technology, instead of photogrammetric method, active sensors (ground based and or airborne based) are being used widely, to directly measure the depth of objects, which provide an ideal data set for city modeling. Furthermore, in hybrid method, large-scale city modeling systems receive most of their data from ground and aerial image sensors, aerial active sensors, and 2D footprint data from GIS or CAD data. Each of these data sources and their corresponding modeling techniques has advantages and disadvantages. For example, images provide texture and color information with high accuracy, making them necessary for texture data and appealing for extracting small model features. Sensors provide accurate footprint and height data whereas ground sensors provide accurate facade data [5,6]. Fusing these data sources can generate more accurate and automatic urban models.

However, in many cases, sensors are too expensive whereas skilled person is also needed to handle them sufficiently. Especially, for developing countries where research cost is limited and adequate person to competently handle sensors is rare, low-cost and simpler method is desirable. Also, as mentioned above, when physical rough factors of the structures (e.g. height of buildings, range between buildings, etc) are thought of being more essential than reality and aesthetic aspects (for example, in this case of a 3D city model reconstructed for the use in traffic and air pollution simulation), there is no urgent need to obtain complex details of the structures. Considering that height of structures itself can be acquired by analyzing shadow of the structures, it is possible to construct urban models with a monocular aerial image as the data source.

Shadow plays an essential role in the scene interpretation by humans, because it gives hints on the third dimension. Even if we have only one monocular view, we are able to derive information about the shape of an object – simply by analyzing its shadow. The idea to derive 3D information from shadow is not new [7] and quite a number of similar and comparable approaches can be found [8,9]. However, the characteristic feature of this paper is that we put emphasis on the optimal simplification on shadow extraction and measurement method, for the sake of a low cost automatic city model reconstruction system.

AUTOMATIC RECONSTRUCTION OF CITY MODEL USING SHADOW ANALYSIS OF AERIAL IMAGERY

THE PRINCIPLE OF SHADOW ANALYSIS

There are three primary assumptions adopted here:

1) The objects being measured are vertical
2) The shadows are cast from the top of the object and not the sides
3) The shadows fall on open, level ground

Under these assumptions, height (h) of the object being measured can be derived from a trigonometrical formula with shadow length (s) and solar zenith (θ) as its variables.
(Fig.1, Formula 1)

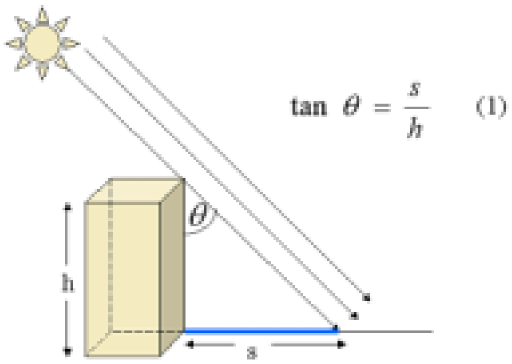


Fig.1. Deriving object's height from shadow length

To determine shadow length, knowledge on solar azimuth direction is needed. Considering that each building has more than one edge whereas two of the edges direct towards solar azimuth direction, solar azimuth can be derived easily by analyzing the direction of the edges of buildings.

THE PRINCIPLE OF IMAGE BINARIZATION

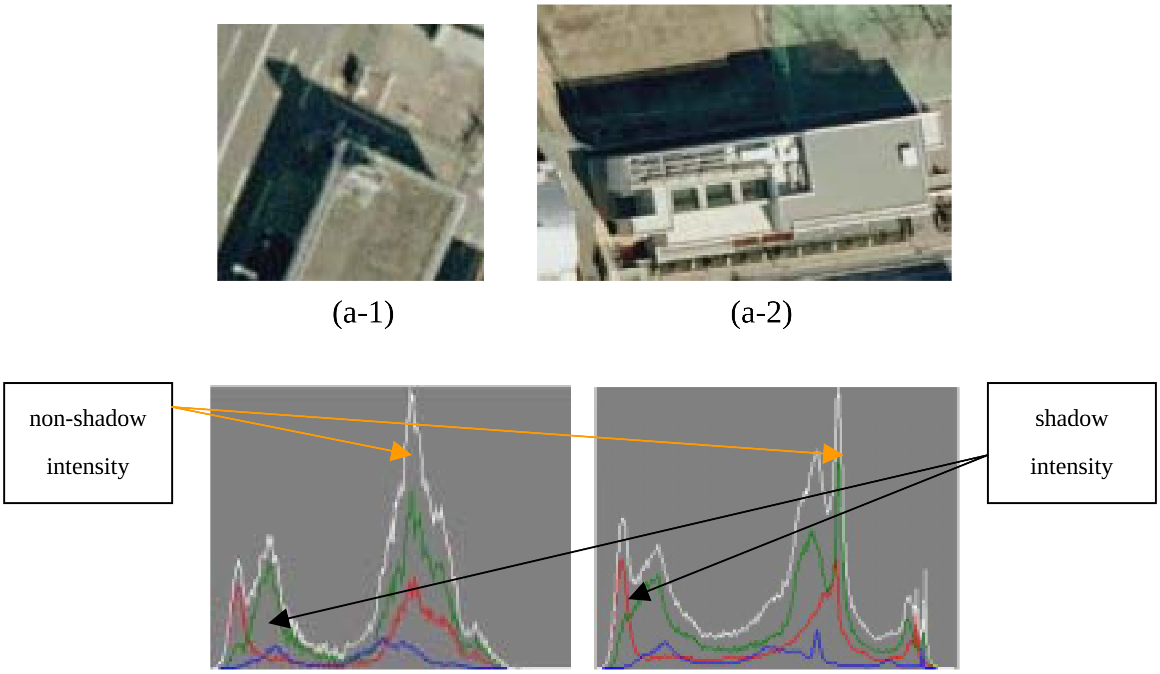
When a value of color intensity is submitted to be the threshold t, in order to binarize an image consists of pixels (x, y) each with color intensity f(x, y), the binary image should consist of pixels (x, y) each with color intensity b(x, y) as follows: [10]

$$b(x,y) = \begin{cases} 1 & (f(x,y) \geq t) \\ 0 & (f(x,y) < t) \end{cases} \qquad (2)$$

Hence, a binary image consists of pixels with two color intensity value only: 0 or 1.

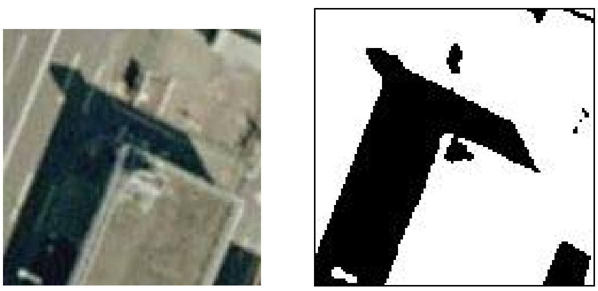
SHADOW EXTRACTION BY IMAGE BINARIZATION

We found that all aerial images have a unique characteristic that the color intensity of shadow area is much lower than non-shadow area (Fig.2). This characteristic then allows the binarization of aerial images to separate shadow area from non-shadow area, under a threshold value which is a mid-value between the intensity of shadow and non-shadow area. (Fig.3)



(b-1) (b-2)

Fig.2. A unique characteristic of aerial images:
non-shadow area color intensity >> shadow area color intensity



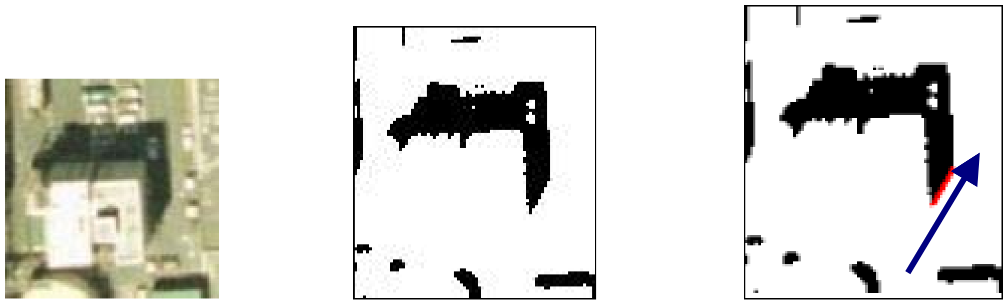
(a) Original Image (b) Binary Image (t=80)

Fig.3. Aerial Image Binarization

MEASURING SHADOW LENGTH USING SOLAR AZIMUTH VECTOR

When solar azimuth direction is known, shadow length of each building can be easily derived by generating solar azimuth vector which then detects the edges of each building's shadow area that are parallel to it. In this process, the following algorithm is applied:

1) Search for the edges which are parallel to solar azimuth vector. If there is no fatal noise on the parallel edges, then two parallel edges will be obtained. However, it is difficult to obtain two parallel edges with exactly the same length, due to image noise, etc. In such a case, the length of shadow is the length of the shorter one. (Fig.4)
2) However, if there is no parallel edge obtained, multiple lines parallel to solar azimuth vector are generated within shadow area so that multiple parallel lines obtained. Calculating the average of the length of first longest n parallel lines then derives the length of shadow. (Fig.5)



(a) Original Image (b) Binary Image (t=80) (c) An edge parallel to Solar Azimuth Vector

Fig.4. Searching for the edges which are parallel to solar azimuth vector.

(a) Original Image (b) Binary Image (t=80) (c) Multiple Parallel Lines

Fig.5. Generating multiple parallel lines for deriving shadow length

GENERATING GEOMETRY MODEL

(a) 2D map (b) 3D city model

Fig.6. Adding the 3rd dimension (height) to 2D map results in 3D city model

When shadow length is known, object's height can be derived easily. No need to calculate solar zenith since solar zenith is defined unanimously for every single urban area. Hence, based on above Formula 1, the following equation is obtained:

$$\tan\theta = \frac{s_1}{h_1} = \frac{s_2}{h_2} = \ldots = \frac{s_n}{h_n} \qquad (3)$$

Thus, instead of calculating solar zenith, which must be derived using so many variables that we should have enough knowledge on photograph condition (e.g. photograph date and time, airplane flying height, etc), it is desirable that there is a structure in the measured image with its height is known previously.

When a ground plan of structures (e.g. 2D digital map) is available, 3D geometry models can be generated by adding the obtained height data as the third dimension information. (Fig.6)

DYNAMIC TRAFFIC AND AIR POLLUTION SIMULATION

In this work, vehicle position is given as real world coordinates (longitude and latitude). These real world coordinates are then translated into virtual world ones. Every single vehicle releases air pollutants that disperse into the air obeying the rules of fluid dynamics.

As mentioned above, air pollutants quantity released by vehicles is parallel to traffic quantity whereas traffic quantity itself dynamically changes time to time. Hence, time is adopted to be the 4th dimension. Released pollutants are limited to CO and NOx for now.

This work is now in progress under a joint research with Kuwahara Laboratory and Kato-Ooka Laboratory of Institute of Industrial Science, The University of Tokyo.

CONCLUSION

An automatic approach of the reconstruction of large-scale city model based on shadow analysis of aerial imagery is proposed. As an example of the application area of the obtained virtual environment, a dynamic 4-dimensional traffic and air pollution simulation system is introduced. Since this work is still in progress, some problems, including error handling due to assumptions adopted, will be discussed in future.

REFERENCES

(1) J. Hu, S. You, and U. Neumann, “Approach to Large-Scale Urban Modeling”, IEEE Computer Graphics and Applications, vol.27, no.2, 2003, pp.62-69.

(2) C. Lin and R. Nevatia, “Building Detection and Description from a Single Intensity Image”, Proc. Computer Vision and Image Understanding (CVIU), vol. 72, no. 2, Elsevier Science, 1998, pp. 101-121.

(3) H. Vanden and A. Frank, “3D Reconstruction from a Single Image Using Geometric Constraints”, PandRS, vol. 53, no. 6, Dec. 1998, pp. 354-368.

(4) C. Baillard and H. Maître, “3D Reconstruction of Urban Scenes from Aerial Stereo Imagery: A Focusing Strategy”, Computer Vision and Image Understanding, vol. 76, no. 3, Dec. 1999, pp. 244-258.

(5) C. Früh and A. Zakhor, “Constructing 3D City Models by Merging Ground-Based and Airborne Views”, Proc. Conf. Computer Vision and Pattern Recognition, IEEE CS Press, 2003, pp. 562-569.

(6) V. George, “Fusion of Laser Scanning Data, Maps and Aerial Photographs for Building Reconstruction”, Proc. Int’l Geoscience and Remote Sensing Symp. (IGARSS), IEEE Press, 2002.

(7) M. Sudhakar, A. Girjapurkar, and K.S.S. Subramanian, “Creation of Digital City Model from a Nadir Aerial Photo Using Building Shadow”, RMSI Map India, 2004.

(8) S. Hinz, F. Kurz, A. Baumgartner, and P. Wasmeier, “The Role of Shadow for 3D Object Reconstruction from Monocular Images”, Optical 3D Measurement Techniques (V), Institute of Geodesy and Geophysics, Vienna Univ. of Technology, 2001.

(9) R.B. Irvin and D.M. McKeown, “Methods for Exploiting the Relationship Between Buildings and Their Shadows in Aerial Imagery”, IEEE Transactions on Systems, Man, and Cybernetics 19(6), 1989, pp.1564-1575.

(10) T. Agui and T. Nagao, “Image Processing and Recognition”, Shoukoudou, 1996.