3D ROAD EXTRACTION USING A STEREO PAIR OF AERIAL IMAGES

A. P. Dal Poz *, J. F. C. Silva, R. A. B. Gallis

Dept. of Cartography, São Paulo State University, R. Roberto Simonsen, 305, Presidente Prudente-SP, Brazil, {aluir, fjcsilva}@fct.unesp.br, rodrigogallis@gmail.com

Earth Observation for a Changing World

KEY WORDS: Dynamic Programming, Road Model, Feature Extraction, Road Extraction, Aerial Images

ABSTRACT:

Road extraction from aerial and satellite imagery is of fundamental importance in the context of spatial data capturing and updating for GIS (Geographic Information System) applications. This article proposes a semiautomatic methodology for 3D road extraction from digital images. The dynamic programming (DP) algorithm is used to carry out the optimisation process in the object-space, instead of doing it in the image-space such as the DP traditional methodologies. This means that road centrelines are directly traced in the object-space, implying that a mathematical relationship is necessary to connect road points in object- and image-space. This allows the integration of radiometric information from images into the associate energy function. Although the methodology can operate in different modes (mono-plotting or stereo-plotting), and with several image types, including multisensor images, this paper presents details of our stereo-based approach involving a stereo pair of aerial images. In order to initialise the extraction procedure, a few seed points are necessary to be supplied in the object-space to sparsely and coarsely describe the roads. We visually identify these points on the stereo images and project them onto the object-space. Each road is extracted separately by an iterative scheme based on the DP optimisation algorithm. Each iteration cycle is carried out by means of a polyline densification followed by a DP optimisation. The polyline densification is carried out by linearly interpolating middle points between every adjacent vertices of the current polyline. Preliminary results have shown that the proposed methodology works properly.

1. INTRODUCTION

Most of the studies on the subject of road extraction have been based on semiautomatic or automated strategies formulated in the image-space. Most common semiautomatic approaches are road-follower (e.g.: McKeown and Denlinger, 1988; Kim et al., 2004) and snakes-based methods (e.g.: Kass et al., 1987; Grüen and Li, 1997; Hu et al., 2004). An example of automated approach is found in Baumgartner et al. (1999), in which different resolutions, grouping, and context are used to extract road networks from high-resolution images. Another example is found in Stoica et al. (2004), in which the road network in remote sensed images is modelled as connected line segments, resulting in a probabilistic model to be solved by the Maximum a Posteriori (MAP) estimation.

Very few approaches have been developed to directly track roads in the object-space (Grüen and Li, 1997, Zhang, 2004, Hinz and Baumgartner, 2003). These approaches may be based on an image (Dal Poz et al., 2006) combined with a Digital Terrain Model (DTM) or on two or more images (Grüen and Li, 1997). The main advantage of tracking roads in the object-space is that object-space constraints, like the smoothness vertical control, can be easily introduced into the road extraction strategy. This kind of constraint cannot be enforced when using image-space approaches.

This article proposes a 3D road extraction methodology from a stereo pair of aerial images based on the dynamic programming (DP) optimisation. It is an extension of Dal Poz et al. (2006), in which the single image mode was discussed. This paper is organized in four main sections. Section 2 presents the basic principles of our object-space road model and our strategy for optimising it based on the DP algorithm. Preliminary results are presented and discussed in Section 3. Conclusions are provided in Section 4.

2. METHODOLOGY

2.1 Overview

Road extraction in the object-space is a process that allows the direct extraction of roads in a 3D object-space reference system. Our concept is based on a 3D road model, which is formulated in the object-space and is optimised by DP optimisation. In order to formulate a road model in the object-space in such way that image radiometric information can also be taken into account, it is necessary to establish a rigorous mathematical relation between the object-space and the image-space.

As this road extraction strategy is semiautomatic, an initialisation step of seeding a few points that coarsely describe the road is necessary to start the extraction process. Different strategies may be used to supply the necessary seed points. For example, an operator may detect roads in an image and measure a few seed points along every road. Since road extraction is accomplished in the object-space, the seed points need to be transformed into the object-space reference system.

The object-space road extraction methodology may work in two modes: single image and multiple images. In the single image mode, an aerial image is used to formulate the road model, but a DTM is also necessary. Other kinds of images, such as pushbroom images, can be used if appropriate mathematical models relating the image and object reference systems are available. In the multiple images mode, two or more aerial images can be used to formulate the road model. In this mode,

* Corresponding author.

different sensor images can be combined, leading to multisensor approaches. In both modes, the DP optimisation would generate polylines in the object-space reference system.

2.2 Road model based on a stereo pair of aerial images

Photometric and geometric road properties can be used to formulate in image-space a generic road model considering that the road *r* (Figure 1) can be represented by a polyline $P^i = \{p_1, ..., p_n\}$, where p_i is its ith vertex. Examples of road properties are: roads are elongated and lighter than the background; road grey levels do not change much within a short distance; roads are smooth curves; road centreline points are middle points of road edge points existing along road cross sections. Taking into account these properties and the collinearity-based relationship between a road image point (p_i) and its homologous (P_i) on the object-space road (*R*) (Figure 1), Dal Poz et al. (2006) formulated a road model in object-space that can be generally expressed as,

$$E = \sum_{i=1}^{n-1} E_i(P_{i-1}(E_{i-1}, N_{i-1}, h_{i-1}), P_i(E_i, N_i, h_i), P_{i+1}(E_{i+1}, N_{i+1}, h_{i+1}), W_{i-1}, W_i, W_{i+1}) \quad (1)$$

where: E is the total energy equation, which is decomposed in a sum of n-1 sub-functions E_i

P_{i-1}, P_i, and P_{i+1} are consecutive points of the object-space Polyline $P^o = \{P_1, ..., P_n\}$ representing a road centreline

W_{i-1}, W_i, and W_{i+1} are road widths at points P_{i-1}, P_i, P_{i+1}, respectively

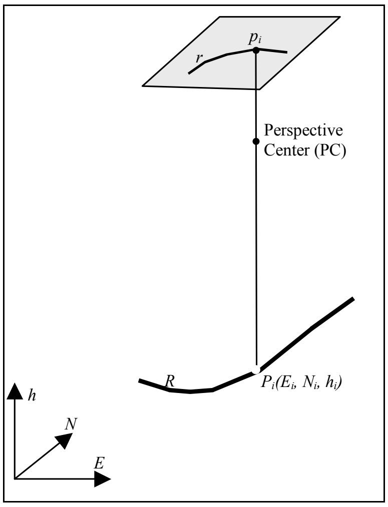

Figure 1. Object-space road model principle for a single aerial image

The ground reference coordinate system used to represent the 3D vertices of road polylines is the UTM (Universe Transverse Mercator) co-ordinates (*E, N*) plus the ellipsoidal height (*h*) (Wolf and Dewitt, 2000).

Equation 1 is ambiguous as it can theoretically assume the same value for an unlimited set of object-space polylines. This results from the well-known property of the collinearity-based relationship between image and object road points, in which one can select infinite object points that map to the same image point. Equation 1 assumes the maximal value for the object-space polyline that represents the road centreline. However, the same occurs for an unlimited set of object-space polylines. Consequently, Equation 1 cannot be used for extracting roads without imposing constraints to remove its ambiguity. Dal Poz et al. (2006) uses a single aerial image combined with an underlying DTM to remove the ambiguity of Equation 1 and to allow the extraction of roads in object-space.

Another way to remove the ambiguity of equation is illustrated in Figure 1. Entities r and r′ are two homologous roads in the left and right images, respectively. Entity R is the corresponding road in the object-space. Let E^l and E^r be the energy equation for the left and right images, respectively. The question here is how to combine both energy equations in order to get another one, unambiguous energy equation. It is possible to prove that the energy function given by the sum of both energy functions E^l and E^r is not ambiguous, i.e., it has a unique solution for the road centreline of road R.

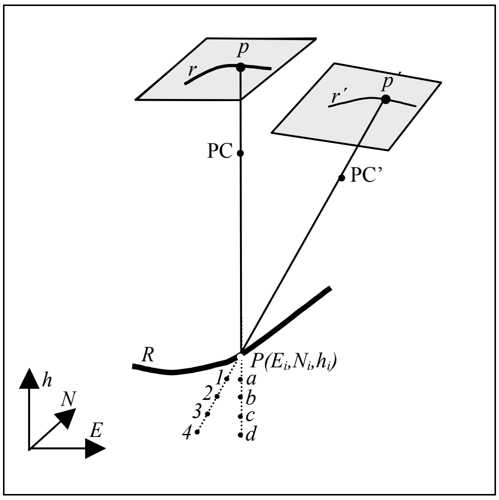

Figure 2. Object-space road model principle for a stereo pair of aerial images.

If the polyline $P^o = \{P_1(E_1, N_1, h_1); P_i(E_i, N_i, h_i); \ldots ;P_n(E_n, N_n, h_n)\}$ correctly describes the road centreline of road R (Figure 2), then it corresponds to the maximum of both energy functions E^l and E^r. The energy equation E^l takes the same maximum if point $P_i(E_i, N_i, h_i)$ were replaced by either point *a* or point *b* or point *c* or point *d* (Figure 2), but the energy equation E^r would take the maximum only for point $P_i(E_i, N_i, h_i)$. The reverse would occurs if point $P_i(E_i, N_i, h_i)$ were replaced by either point *1* or point *2* or point *3* or point *4* (Figure 2). This means that only the polyline P^o corresponds to the maximum of both energy equations and, since they are also positive functions, the energy equation (E^{st}) given by the sum of E^l and E^r takes the maximum only for the polyline P^o. Another advantage of this definition is that it preserves the structure of Equation 1, i.e., the energy equation E^{st} interrelates simultaneously only twelve variables (coordinates of points P_{i-1}, P_i, and P_{i+1} and road widths at these points). As such, equation E^{st} can be solved by an optimisation procedure by the DP algorithm. Mathematical foundations and algorithm aspects of DP optimisation can be found in an extensive literature, e.g. in Ballard and Brown (1982).

2.3 Optimisation strategy by DP algorithm

Our strategy to optimise the energy equation E^{st} starts with an initial 3D polyline coarsely representing a road, which is progressively refined during the progress of the DP optimisation. Polyline vertices are usually referred to as seed points and may be measured on the stereo pair of images and photogrammetrically transformed into the object-space. The initial polyline is refined based on an iterative scheme.

Each iteration cycle is carried out by means of a polyline densification followed by a DP optimisation. The polyline densification is carried out by linearly interpolating middle points between every adjacent vertices of the current polyline. After the densification, the resulting polyline is used as reference for generating a search space composed by polylines that are candidates for the optimal polyline at the current iteration. Candidate polylines can be generated during the DP optimisation by leaving every vertex of the reference polyline to move around its initial position. To avoid search windows with large number of elements (vertices) and to assure a large pull-in-range, we use a planar-like search window, which consists in a regular grid of points sampled within the plane perpendicular to the current polyline at each vertex. Another advantage of using a planar-like search window is that the coordinates of each vertex in the planar search space are mathematically interrelated by the plane equation. As a result, only six independent coordinates of polyline vertices are simultaneously interrelated in the energy function E^{st}. In addition, since road width does not change much along short road segments, we use only one road width per road section defined by three consecutive points. Thus, the final energy equation E^{st} interrelates simultaneously seven variables.

The optimisation process proceeds until a convergence criterion is satisfied. Convergence checking consists in verifying after each iteration cycle whether all added points are collinear to neighbour points.

3. EXPERIMENTAL RESULTS

The proposed methodology was implemented using Borland C++ Builder 5 compiler for Windows XP. One stereo pair of high-resolution aerial images (9286 x 9496 pixels) at the approximate scale of 1:9200 is used in experiments, in which the pixel footprint is about 25 cm. These images are from a region of Switzerland and are available in the LPS (Leica Photogrammetry Suite®) system, along with interior and exterior orientation parameters. Below we present two examples of results showed in two stereo pairs of sub-images. These stereo sub-images were selected from the processed stereo aerial images. A qualitative analysis of the results is provided by visually inspecting the extracted roads overlaid on the stereo sub-images.

Figure 3 shows one example of extraction of a road segment from a stereo pair of sub-images. This road segment shows a good contrast with the background and no important anomalies are presented along it. Although the road centreline overlaid on the top sub-image is slightly better, both overlaid road centrelines can be considered accurate. Slight different geometric qualities of both road centrelines overlaid on the top and bottom sub-images can be attributed to the fact that their projections are carried out by two different image orientation parameter sets. In general, it is possible to note that the extracted road centreline overlaid on the top sub-image is slightly closer to left road margin and the opposite occurs with the road centreline overlaid on the bottom sub-image.

Figure 3. Extraction of a road segment from a stereo sub-images.

Figure 4 shows another example of extraction of road segments from a stereo pair of sub-images. The contrast of the road with the background is good, but some minor anomalies are presented along the longer road, as e.g. perspective obstruction caused by trees, shadows cast by trees, and missing edge around the road intersection. Only a slight direction change of the overlaid road centrelines can be observed along the road intersection. As in previous example, probably due to differences in quality of both exterior image orientation parameter sets, the projection of roads onto the image-space is more accurate for the bottom sub-image.

Figure 4. Extraction of road segments from stereo sub-images.

4. CONCLUSIONS

In this paper an object-space road extraction methodology from a stereo pair of aerial images was proposed. In order to initialise the extraction procedure, a few seed points are necessary to be supplied in the object-space. We identify these points on the stereo images and project them onto the object-space.
In order to preliminarily exemplify the performance of the proposed approach, we presented and analysed two examples of results showed in two pairs of stereo sub-images selected from the input stereo pair of aerial images. First example showed a well-contrasted road segment. Second example showed two concurrent roads and the longer one is affected by some anomalies, as e.g. perspective obstruction caused by trees, shadows cast by trees, and a missing edge at the road intersection. Road centrelines overlaid on the stereo pairs of sub-images showed that the results for both cases are relatively accurate, except for small projection errors.

REFERENCES

Ballard, D. H.; Brown, C. M., 1982. *Computer Vision*. Prentice Hall, Inc., Englewood Cliffs, New Jersey, 523p.

Baumgartner, A., Steger, C., Mayer, H., Eckstein, W. Ebner, H., 1999. Automatic Road Extraction Based on Multi-Scale, Grouping, and Context. *Photogrammetric Engineering and Remote Sensing*, 65(7), pp. 777–785.

Dal Poz, A. P., Vale, G. M., 2003. Dynamic programming approach for semi-automated road extraction from medium- and high-resolution images. In: *ISPRS Commission III Symposium – PIA´03*. Munich, Germany, pp. 87-92.

Dal Poz, A. P., Gallis, R. B. A., Silva, J. F. C., 2006. Semiautomatic road extraction by dynamic programming optimisation in the object space: Single image case. In: *ISPRS Commission III Symposium – PCV´06*, Munich, Germany, pp. 215-220.

Grüen, A., Li, H., 1997. Semi-automatic linear feature extraction by dynamic programming and LSB-snakes. *Photogrammetric Engineering and Remote Sensing*, 63(8), pp 985-995.

Hinz, S., Baumgartner, A., 2003. Automatic extraction of urban road networks from multi-view aerial imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(1-2), pp. 83-98.

Hu, X., Tao, C.V., Hu, Y., 2004. Automatic Road Extraction from Dense Urban Area by Integrated Processing of High Resolution Imagery and Lidar Data. In: *XXth ISPRS Congress*, Istanbul, Turkey. (CD-ROM).

Kass, M., Witkin, A., Terzopoulos, D., 1987. Snakes: Active Contour Models. In: *1st International Conference on Computer Vision*, London, England, pp. 259-268.

Kim, T., Park, S-R., Kim, M-G., Jung, S., Kim, K-O., 2004. Tracking road centerlines from high resolution remote sensing images by least squares correlation matching. *Photogrammetric Engineering and Remote Sensing* , 70(12), pp. 1417-1422.

McKeown, D. M., Delinger, J. L., 1988. Cooperative methods for road tracking in aerial imagery. In.: Workshop of Computer Vision and Pattern Recognition, pp. 662-672.

Stoica, R., Descombes, X., Zerubia, J., 2004. *A* Gibbs Point Process for Road Extraction from Remotely Sensed Images. *International Journal of Computer Vision*, 57(2), pp. 121-136.

Wolf, P. R., Dewitt, B. A., 2000. *Elements of Photogrammetry with Applications in GIS*. McGraw-Hill, 3rd Ed, Boston, MA, USA, 589p..

Zhang, C., 2004. Towards an operational system for automated updating of road databases by integration of imagery and geodata. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(3-4), pp. 166-186.

ACKNOWLEDGEMENTS

This work was supported by FAPESP (Research Foundation of the State of São Paulo, Brazil) and CAPES (Brazilian Foundation for the Support of Graduate Studies).