THREE-DIMENSIONAL ROAD EXTRACTION COMBINING A STEREOSCOPIC PAIR OF LOW-RESOLUTION AERIAL IMAGES AND A DTM

A. P. Dal Poz a,*, E. F. O. Martins b, R. B. Zanin b

a Dept. of Cartography, São Paulo State University, Presidente Prudente-São Paulo, Brazil - aluir@fct.unesp.br
b Dept. of Mathematics, Mato Grosso State University, Sinop-MT, Brazil - (profericomartins, rodrigo.zanin)@unemat-net.br

KEY WORDS: Dynamic programming; road model; feature extraction; road extraction; aerial images

ABSTRACT:

This paper proposes a semiautomatic method for road extraction. The proposed method combines a stereoscopic pair of low-resolution aerial images with a polyhedron generated with a digital terrain model (DTM). The problem is formulated in object space by means of an objective function that models the object ‘road’ as a smooth curve belonging to a polyhedral surface. The proposed objective function depends on radiometric information accessed in the image space via a collinear relationship between road points in the object space and corresponding points in the image spaces of stereoscopic images. The polyline that provides the best representation of a selected road is obtained via the optimization of the objective function using a dynamic programming algorithm. The optimization process is iterative. An operator is required to supply an initial polyline approximating the selected road. The obtained results show that the proposed method is robust even when faced with anomalies along roads, such as obstructions caused by shadows and trees.

1. INTRODUCTION

Methods for road extraction from aerial and satellite imagery are essential for capturing and updating the spatial information. Several studies have been conducted on this subject beginning with the pioneering studies of Bajcsy and Tavakoli (1976) and Quam (1978). The majority of the studies of road extraction have involved models and approaches formulated in the image space. Most common semiautomatic strategies are road trackers (McKeown et al., 1988, Kim et al., 2004) and curve-fitting methods (Kass et al., 1987, Neumenschwader et al., 1997, Agouris et al., 2000, Göpfert et al., 2011). Conversely, automated methods of road extraction require a highly sophisticated integration of context-based information and prior knowledge (Baumgartner et al., 1999; Hu et al., 2007, Poullis e You, 2010). Currently, there are few existing object-space methods (e.g., Grüen e Li, 1997 and Dal Poz et al., 2010, 2012) for road extraction. These previously developed methods are based on a single image (mono mode) combined with a digital terrain model (DTM) or utilize two or more images (stereo mode) from one or more sensors. The stereo mode method can integrate a DTM into the extraction, which can render a more stable representation of a selected road. In these equations, $\alpha$ is a function that depends on the vertex $p_i$ and expresses road pixels as being lighter than the their neighbouring pixels; $E_{pj}$ is a function that depends on the vertex $p_i$ and expresses road pixels as being lighter than their neighbouring pixels; $E_{pj}$ is a function that depends on two

2. METHOD

The proposed method is based on an object space road model formulated by Dal Poz et al. (2012) for a stereoscopic pair of aerial images. Subsection 2.1 provides a brief description of the pre-existing object space road model. Subsection 2.2 describes the modified model and the search-space sampling method.

2.1 Road model based on stereoscopic aerial images

Assuming that a road in a low-resolution image can be represented by the polyline $P = \{p_1, \ldots, p_n\}$, where $p_i$ is the $ith$ vertex, one can generate a mathematical model using an objective function (Equation 1) and the inequality constraint (Equation 2) as follows (Grüen e Li, 1997):

$$E = \sum_{i=1}^{n-1} \left[ E_{p1} - \beta E_{p2} + \gamma E_{p3} \right] \times \left( 1 + \cos(\alpha_i - \alpha_{i+1}) \right) |\Delta_s| \right)$$ (1)

$$C_i = a_i - a_{i+1} < T$$ (2)

In these equations, $E_{p1}$ is a function that depends on the vertex $p_i$ and expresses road pixels as being lighter than their neighbouring pixels; $E_{p2}$ is a function that depends on two

* Corresponding author.
consecutive points \((p_{i-1}, p_i, p_{i+1})\) of polyline \(P\) and specifies that grey or colour levels of roads typically do not vary within short distances; \(E_{p3}\) is a function that depends on vertex \(p_i\) and expresses that the road is a lighter linear feature; \(\alpha_i\) is the direction of the linear segment defined by points \(p_{i-1}\) and \(p_i\); \(\beta\) and \(\gamma\) are positive constants; \(|\Delta s_i|\) is the distance between points \(p_{i-1}\) and \(p_i\); and \(T\) is an angular threshold that limits the change of direction between two successive segments of polyline \(P\).

According to Equation 1, only three consecutive points \((p_{i-1}, p_i, p_{i+1})\) of polyline \(P\) are simultaneously interrelated, and thus, this relationship can be decomposed into a sum of \(n-1\) subfunctions \(E(p_{i-1}, p_i, p_{i+1})\), as shown in Equation 3:

\[
E = \sum_{i=1}^{n} E_i(p_{i-1}, p_i, p_{i+1}) \tag{3}
\]

The model solution is a polyline \(P = \{p_1, \ldots, p_n\}\) that represents a road and corresponds to the maximum value of the objective function given by Equation 3. In object space, this objective function can be formulated in terms of vertex point coordinates that belong to the corresponding polyline. Thus, using the collinearity equation one can establish the mathematical relationship between the line \((L_i)\) and column \((C_i)\) coordinates of vertex \(p_i\) and the object space coordinates of \(P\). By specifying \(P\) using the ellipsoidal height \(h_i\) and defining the coordinates \(E_i\) and \(N_i\) in the Universal Transverse Mercator (UTM) system, Equation 3 can be expressed as:

\[
E = \sum_{i=1}^{n} E_i(p_{i-1}, N_i, h_i) \tag{4}
\]

The objective function expressed in Equation 4 is ambiguous because there is an infinite number of polylines in object space that become a single polyline when projected into image space. To remove this ambiguity and obtain a single solution, Dal Poz et al. (2012) developed an objective function specifically for an aerial imagery stereoscopic pair, which is defined as the sum of the objective functions for the left- and right-hand images (both low resolution):

\[
E' = E'_l + E'_r = \sum_{i=1}^{n} E_i^T(p_{i-1}, N_{i-1}, h_{i-1}) + P_{i-1}^T(E_i, N_i, h_i), P_{i+1}, (E_{i+1}, N_{i+1}, h_{i+1}) \tag{5}
\]

In Equation 5 (Figure 1), \(E'_l\) is the objective function (Equation 4) that correlates road \(R\) in the object space with road \(r\) in the left-hand image; \(E'_r\) is the objective function (Equation 4) that correlates the road centrelines \(R\) with the road centrelines \(r'\) in the right-hand image; and \(E_i^T\) is obtained by grouping like terms of \(E'_l\) and \(E'_r\). The mathematical model presented in Equation 5 is designated the stereoroad model.

Figure 1. Road model principle for a pair of stereoscopic images
(Dal Poz et al. 2012)

2.2 Proposed road extraction method

The extraction process is initiated by defining a polyline in object space from several seed points that are provided by the operator and sparsely placed along the road that will be extracted. As a general rule, a small number of seed points is required along segments with a low curvature, whereas segments with a large curvature will require a greater number of seed points. Because the operator must visualize the road to provide seed points, the adopted strategy measures seed points in one image of the stereoscopic pair and projects them over the DTM using a monorestitution algorithm (Makarovik, 1973).

Figure 2. (a) Initial polyline with three seed points. (b) Densification
The initial polyline (Figure 2a) is then gradually densified and refined over iterative optimization cycles in object space until it adequately describes the road centreline. Densification is performed by linearly interpolating mid-points between each pair of pre-existing adjacent vertices (Figure 2b).

Figure 3. (a) Planes that are perpendicular to the densified polyline. (b) Search polylines defined by intersecting the planes with the DTM. (c) Sampled search polylines

The vertices resulting from the polyline densification are used as references to sample candidates for the optimal vertex. If \( m \) candidate vertices are tested for each of the \( n \) vertices of the densified polyline, then \( mn \) candidates for the optimal polyline will be generated in the search-space at each iteration. It is therefore recommended that the lowest possible number of candidates for each vertex is tested. Accordingly, the search-space is limited to a search polyline (Figure 3b) obtained by intersecting the DTM-generated polyhedron with planes that are perpendicular to the densified polyline (Figure 3a). Candidate vertices are symmetrically and regularly sampled along the search polylines (Figure 3c). The extension of the search polylines is dependent on the spatial proximity of the operator-defined seed points. To maximize the convergence radius and guarantee an accurate result, a multi-resolution strategy is adopted whereby lower-resolution polylines are employed in initial iterations (with a width of the same order as the road, thereby ensuring a wider search region) and a higher resolution is adopted in later iterations (ca. 1/3 of the road width). This process enables the correct positioning of the optimal polyline upon the road.

In the example shown in Figure 3c, eight vertices (yellow dots) are sampled symmetrically at regular intervals along the search polyline. Each central vertex (red dots) corresponds to a vertex of the densified polyline (Figure 3b). Vertex sampling along each search polyline is performed using the 3-D parametric form of the linear equation. Each vertex \( P(E_i, N_i, h_i) \) is computed as a function of the line parameter \( t_i \), which describes the distance between the vertex being sampled \( P(E_i, N_i, h_i) \) and the central vertex. Therefore, because each vertex \( P(E_i, N_i, h) \) will depend only on the distance \( t_i \), Equation 5 can be rewritten as:

\[
E_T = \sum_{i=1}^{n-1} E_T^T \begin{pmatrix} P_i(t_i-1), P_i(t_i), P_{i+1}(t_{i+1}) \end{pmatrix}
\]  (6)

According to Equation 6, only three variables \( (t_{i-1}, t_i, t_{i+1}) \) are simultaneously interrelated as opposed to nine variables as in Equation 5. This structure enables us to use a “time-delayed” discrete dynamic programming algorithm (Ballard & Brown, 1982; Amini et al., 1990) because the underlying condition of this optimization algorithm — few variables are simultaneously interrelated — is fully satisfied.

3. EXPERIMENTAL RESULTS

In the following we present preliminary results based on two stereoscopic pairs of low-resolution aerial sub-images. We used a TIN-based DTM with an average resolution of 1 m. The results were evaluated both visually and numerically. The visual evaluation was performed by analysing and overlaying the extracted polylines onto one of the stereoscopic pair of sub-images. The numerical evaluation was performed by calculating the completeness and correctness parameters. Definitions and formulations of these quality indexes can be found in Heipke et al. (1997).

Figure 4. (a) Seed points. (b) Extracted road centreline
Figure 4a shows a road at a high contrast with the surrounding lateral regions and a very smooth geometry. The primary anomalies are its intersection with another road segment and a small group of buildings along its length. These anomalies cause certain small margin segments to be absent and lead to small variations in road width. Figure 4b demonstrates the performance of the proposed method when applied to this low-complexity segment. This road segment was successfully extracted, as indicated by the maximum score (100%) of the completeness index. However, the correctness index was 83%.

The road segment in Figure 5a shows a long curve with two occlusion regions: one is caused by buildings and projected shadows (Figure 5b, smaller rectangle), and the other is caused by a forest patch that largely obstructs a long road segment. Four seed points were chosen, including two placed close together roughly in between the two central seed points to help capture the road signal in the segment obstructed by the forest. The result of this extraction is satisfactory, although a portion of the centreline is coincident with one of the road borders. This placement effect is a consequence of the method, which is most efficient when modelling roads up to 3 pixels wide, which is narrower than the analysed segment. The completeness and correctness parameters were 100% and 77%, respectively.

**4. CONCLUSIONS**

The present study described a semiautomatic method for road extraction using a stereoscopic pair of low-resolution aerial images and a DTM based on DP optimization in object space. The DTM allows the search for the optimal polyline to be restricted along a narrow band that is overlaid upon the model. It also allows for the elimination of candidate polylines that do not meet the vertical smoothness criterion, i.e., those polylines with a vertical deflection angle that violates a user-defined threshold. To preliminarily evaluate the performance of the proposed method, two experiments were designed using two stereoscopic pairs of low-resolution aerial sub-images and a 1-m-resolution DTM. The results were qualitatively and quantitatively analysed (the quantitative assessment was based on completeness and correctness indices). Visually, the extracted polylines were of good geometrical quality, although the correctness parameter fell below 80% in one case.

**REFERENCES**


