SEMI-AUTOMATIC EXTRACTION OF 3D LINES USING A SCALABLE EDGE MODEL AND LEAST-SQUARES TEMPLATE MATCHING

Xiangyun Hu, ERDAS Inc., Norcross, Georgia, United Sates
Zuxun Zhang, School of Remote Sensing and Information Engineering, Wuhan University, China
Jonathan Li, Department of Geography and Environmental Management
University of Waterloo, Ontario, Canada

The extraction of geospatial features from remotely sensed imagery remains the primary means by which to create or update geospatial databases, in particular, for complex urban areas with rapid and frequent changes. This paper presents a novel scalable edge model for rectifying straight lines and spline curves in three-dimensional (3D) object space. It can be used for accurate extraction of 3D linear features and buildings from a stereo pair of aerial images. Least-squares adjustment is employed to iteratively rectify a shifted initial line to the precise 3D position. The adjustment procedure is integrated with a scalable slope edge profile model and an adaptive template generation process, which can be used to locate the edge position by the 'zero-crossing' property of the model. The experimental results using synthesized image and real stereo aerial image pairs show that this approach is effective and has potential advantage.

L'extraction d'éléments géospatiaux à partir d'images obtenues par télédétection demeure la principale façon de créer ou de mettre à jour des bases de données géospatiales, en particulier pour les régions urbaines complexes qui changent souvent et rapidement. Cet article présente un nouveau modèle de contours adaptables pour rectifier les lignes droites et les courbes splines en objets spatiaux tridimensionnels (3-D). Il peut être utilisé pour extraire les éléments linéaires et les bâtiments avec exactitude à l'aide d'un couple stéréo d'images aériennes. On utilise la méthode des moindres carrés pour rectifier itérativement une ligne initiale déplacée dans sa position 3-D exacte. La procédure d'ajustement est intégrée à un modèle de profils de contours en pente adaptables et à un processus de production de gabarit adaptatif, qui peuvent être utilisés pour localiser la position du contour à l'aide de la fonction « croisement zéro » du modèle. Les résultats des expériences effectuées avec des couples d'images numériques et des couples stéréoscopiques de véritables photos aériennes démontrent que cette approche est efficace et possède un avantage potentiel.

Xiangyun Hu
xiangyun.hu@gmail.com

Zuxun Zhang
zxzhang@supresoft.com.cn

Jonathan Li
junli@uwaterloo.ca

Introduction

Reliable and up-to-date information on the rapid and frequent changes of urban environment are critical to various applications ranging from city planning, environmental impact assessment, microclimate simulation and disaster management. However, the application of the conventional classification methods that rely on pixel-based processing methods encounters many difficulties in the extraction of land cove and land-use information from high-resolution, remotely sensed imagery, in particular, for accurate extraction of three-dimensional (3D) features such as 3D buildings. Linear features (e.g., road, object edge, etc.) play an important role in geographical information collection and application. Accurate extraction of linear features from remotely sensed imagery is useful for applications such as data acquisition, data (vector-image) registration and change detection. Linear feature extraction from a single remotely sensed image has been extensively studied. Given initial 'seeds' or the approximate position of a line, the precise shape and position are consequently extracted by the designed algorithm. This approach is termed 'semi-automatic' extraction, which is essential for rectifying an inaccurate linear feature to its exact position. To date, optimization-based methods have been used frequently to extract features by finding optimal 'routes' depending on global geometric and radiometric constraints of the feature. Dynamic programming [*Grün and Li* 1995], active contour models (snakes) [*Trinder and Li* 1995], least-squares template matching (LSTM) [*Grün and Agouris* 1994; *Hu et al*. 2004; *Kim et al*. 2004], and least-squares B-spline snake (LSB-snake), a combined LSTM and snakes method [*Grün and Li* 1997] are essentially optimization-based methods that minimize 'cost' or 'energy' or 'least-squares

error' in order to obtain a globally optimized and reliable result. In their methods, the radiometric model (or constraint) is relatively simple, for example, maximizing the sum of the gradient magnitudes of the edge. An explicit edge model is useful for obtaining a more accurate result. To optical remotely sensed imagery, there is a broad range of edge scales and noisy levels. *Ye et al*. [2005] used a blurred edge model and a least-squared error-based method to detect edges in subpixel accuracy. Their experimental results indicate that the proposed model and method has better results in terms of accuracy and anti-noise capability compared with using the moment-based method [e.g., *Shan and Boon* 2000] and interpolation method [*Steger* 1998; *Steger* 2000].

This paper deals with rectification of three-dimensional (3D) linear features showing intensity gradient (edge) on stereo aerial images.

This paper deals with rectification of three-dimensional (3D) linear features showing intensity gradient (edge) on stereo aerial images. The mentioned optimization-based methods, for example, LSB-snake [*Grün and Li* 1997], can be extended to rectify 3D features, but few of them take into account the rigorous edge model for the purpose of achieving higher accuracy. The model-based fitting using edge information presented by *Lowe* [1991] suggests minimizing the non-linear error function in the image domain. In this fitting model, the perpendicular distance between the projected model line and extracted edge point is minimized. The correct model line to image edge correspondence is found by selecting the one that has the shortest perpendicular distance, which is dependent on the approximate model position. *Vosselman and Veldhuis* [1999] followed the procedure of *Lowe* [1991], with the exception that *Lowe* [1991] uses edge information and a minimization of error function, whereas they use gradient values as weights in observation equations in the least-squares system. *Tseng and Wang* [2003] also employed least-squares adjustment to fit the building wireframe model to the edge points extracted from stereo aerial images. There could be a drawback in this edge point fitting method when the extracted edge pixels are not accurate in geometry due to the varying edge scales and noises. In this paper, instead of using perpendicular distance as the observations in the least-squares adjustment, we integrate a scaleable slope (blurred) edge model with least-squares template matching (LSTM) to accurately rectify linear features from stereo aerial images, overcoming negative effects of varying edge scales and noises. The computational model is derived and experimental results are given, also showing that the potential of the proposed method in computer vision applications as the image formation geometry is essentially the same as the stereo aerial images. The objective of the method is to develop a robust and user-friendly tool that can be integrated into a digitizing production with high reliability, efficiency and accuracy [*Hu et al.* 2004].

Strategy for Semi-automatic 3D Line Extraction

There were two major research directions in which image-based feature extraction strategies and systems were developed: fully automated and semi-automated. A fully automated system, conceptually, produces a scene model in which no user interaction is required. By contrast, a semi-automated system requires user and automated processes in the system to work in sequence. The semi-automated method attempts to integrate the intelligence of our human visual system with an ability to recognize the object robustly and the computer system with an ability to perform fast feature extraction and accurate shape representation. To ensure that the semi-automated extraction can be applied in an operational environment, the method needs to guarantee better performance in terms of (1) reliability: extraction should be non-sensitive to noise such as shadows and occlusions, editing for correction of the extracted results should be minimized; (2) accuracy: the geometric error of the extracted results should be minimized and the extracted results should be at least comparable to that of manual digitizing; (3) efficiency: the extraction time should be much less than that required by manual operation; (4) interactivity: the extraction should be an interactive process in which the human operator is able to correct wrongly extracted results immediately. High efficiency is also dependent on reliability, accuracy and interactivity. Interactivity with timely feedback from the automated process is an important factor influencing high extraction efficiency [*Hu et al.* 2004].

Given the importance of linear features, this paper deals mainly with the algorithm development of semi-automatic 3D line extraction. The main feature of the proposed 3D line extraction strategy is that, based upon photogrammetric geometry, it utilizes a scalable edge model for rectifying straight lines and spline curves in 3D object space. The method can be used for accurate extraction of 3D building roofs from a stereo pair of aerial images. Least-squares template matching (LSTM) is employed to iteratively rectify a shifted initial line to the precise 3D position. The LSTM procedure is integrated with an adaptive template generation process, which can be used to locate the edge position by the 'zero-crossing' property of the model.

LSTM for 3D Linear Feature Rectification

Figure 1 depicts the formation of a pair of stereo aerial images. The object reconstruction from stereo aerial images is based on project geometry expressed by the following collinearity equations [*Wolf and Dewitt* 2000]

$$x_a - x_0 = -f\frac{m_{11}(X_p - X_L) + m_{12}(Y_p - Y_L) + m_{13}(Z_p - Z_L)}{m_{31}(X_p - X_L) + m_{32}(Y_p - Y_L) + m_{33}(Z_p - Z_L)} \quad (1)$$

$$y_a - y_0 = -f\frac{m_{21}(X_p - X_L) + m_{22}(Y_p - Y_L) + m_{23}(Z_p - Z_L)}{m_{31}(X_p - X_L) + m_{32}(Y_p - Y_L) + m_{33}(Z_p - Z_L)} \quad (2)$$

where x_a and y_a are the photo coordinates of image point a; X_p, Y_p, and Z_p are the object space coordinates of the corresponding point a; X_L, Y_L, and Z_L are the object space coordinates of the exposure station; f is the camera focal length; x_0 and y_0 are the coordinates of the principal point (usually known from camera calibration); and the m_{ij}, where i = 1, 2, 3, and j = 1, 2, 3, are the functions of three rotation angles, ω, ϕ, and x. The collinearity equations are nonlinear and can be linearized by using Taylor's theorem. If ω, ϕ, x and X_L, Y_L, Z_L and are known, using the known image coordinates on the stereo pair, we can obtain the object space coordinates, X_p, Y_p, and Z_p of the ground point P. Consequently, the errors of a point position on the image, dx_p and dy_p, are caused by the errors of the point position in object space, dX_p, dY_p, and dZ_p. The linearized error equations are expressed as

$$dx_p = -\frac{\partial f^x}{\partial X}dX_p + \frac{\partial f^x}{\partial Y}dY_p + \frac{\partial f^x}{\partial Z}dZ_p = A_x dX_p + A_y dY_p + A_z dZ_p \quad (3)$$

$$dy_p = -\frac{\partial f^y}{\partial X}dX_p + \frac{\partial f^y}{\partial Y}dY_p + \frac{\partial f^y}{\partial Z}dZ_p = B_x dX_p + B_y dY_p + B_z dZ_p \quad (4)$$

where A_X, A_Y, A_Z, B_X, B_Y, and B_Z are the linearized coefficients based on the Taylor series. The detailed description on the linearization of the collinerity equations can be found in *Wolf and Dewitt* [2000]. In this paper, 3D linear feature rectification depends on equations (3) and (4).

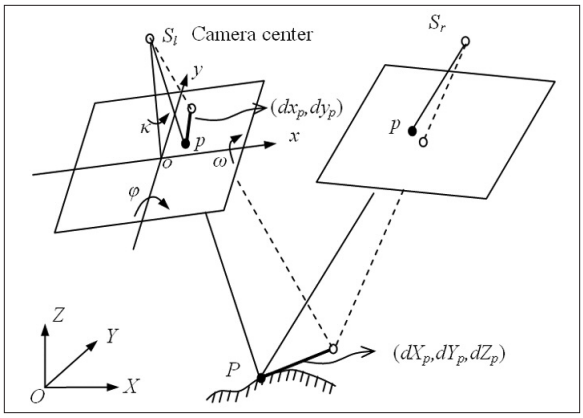


Figure 1: Illustration of 3D restitution of stereo images.

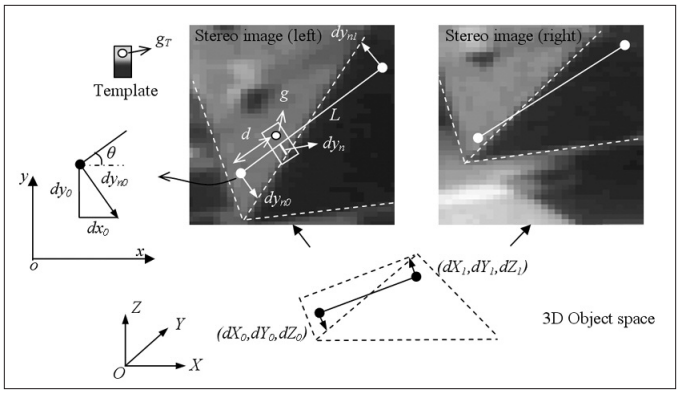


Figure 2: 3D Straight line rectification.

We first derive the computational model to rectify straight lines in 3D, followed by the derivation of the rectification of 3D spline. As shown in Figure 2, the solid line lies in the initial position and needs to be rectified to its exact position indicated by the dashed lines in both image and object space. The length of the line on an image is L. dy_{n0} and dy_{n1} is the shift along the normal direction at the end point of the line. For a pixel p close to the initial line and located in d, let its gray level be g and its position error in the normal direction be dy_n. Assuming the image template indicating the exact local edge model is already known (the gray value on the corresponding pixel is g_T), in order to obtain minimal $\sum_i (g_T - g)^2$, where i indicates the number of a pixel inside the template, we have $g_T = g(y_n + dy_n)$ and the error function

$$v_g = \frac{\partial g}{\partial y_n}\partial y_n - (g_T - g) + g_y dy_n - l_g \quad (5)$$

where g_y is the gray gradient in the normal direction, l_g is the error of the gray value between the pixel and the corresponding template pixel, and dy_n is the shift along the normal direction. From

$$dy_n = (d / L) \cdot dy_{n0} + ((L - d) / L) \cdot dy_{n1} \quad (6)$$

and equation (5), we have

$$(d / L) \cdot g_y\, dy_{n0} + ((L - d) / L) \cdot g_y\, dy_{n1} = l_g \quad (7)$$

As shown in Figure 2, we know

$$dy_{n0} = - dx_0 / \sin\theta \quad (8)$$

or

$$dy_{n0} = -dy_0 / \cos\theta \quad (9)$$

θ is the angle of the line on the image. When $|\sin\theta| > |\cos\theta|$ we use equations (3), (4), (7), and (8) to derive the adjustment model to rectify the 3D line:

$$m_0 \cdot dX_0 + m_1 \cdot dY_0 + m_2 \cdot dZ_0 + m_3 \cdot dX_1 + m_4 \cdot dY_1 + m_5 \cdot dZ_1 = l_g \quad (10)$$

where

$$m_0 = -((d / L) g_y A_x) / \sin\theta$$
$$m_1 = -((d / L) g_y A_y) / \sin\theta$$
$$m_2 = -((d / L) g_y A_z) / \sin\theta$$
$$m_3 = ((d - L) / L)\, g_y A_x) / \sin\theta$$
$$m_4 = ((d - L) / L)\, g_y A_y) / \sin\theta$$
$$m_5 = ((d - L) / L)\, g_y A_z) / \sin\theta$$

when $|\sin\theta| \le |\cos\theta|$, using equations (3), (4), (7), and (9), we can derive a similar model for the rectification. We can iteratively shift the end point of the line with (dX_0, dY_0, dZ_0) and (dX_1, dY_1, dZ_1) until the terminating condition is satisfied.

The interpolation spline, cardinal spline, is used to represent 3D curves. A cardinal spline is a cubic Hermite spline whose tangents are defined by points and a tension parameter. A cardinal spline takes the positions of the current point and, along with the previous and next points, averages the positions using the tension parameter. This smoothes the line and denotes a path that is gently curved through the points rather than zigzagging through them. As shown in Figure 3, a piece of the cardinal spline defined by the four continuous control points P_{i-1}, P_i, P_{i+1}, and P_{i+2} by

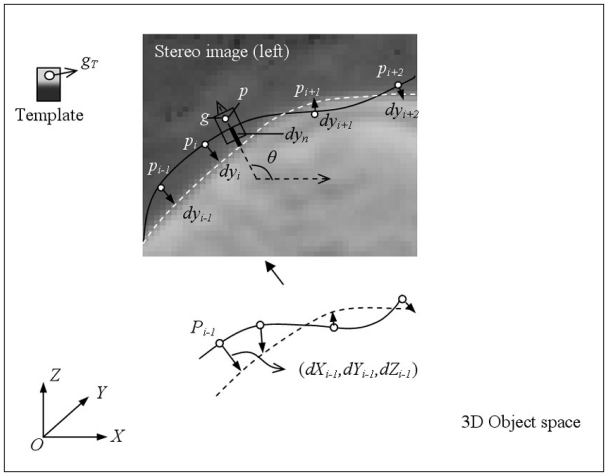


Figure 3: 3D spline rectification.

$$P(u) = P_{i-1}(-su^3 + 2su^2 - su) + P_i[(2 - s)u^3 + (s - 3)u^2 + 1] + P_{i+1}[(s - 2)u^3 + (3 - 2_s)u^2 + su] + P_{i+2}(su^3 - su^2) = P_{i-1}c_0(u) + P_i c_1(u) + P_{i+1}c_2(u) + P_{i+2}c_3(u) \quad (11)$$

where u is the parameter and $0 \le u \le 1$, $s = (1 - t) / 2$, t is the tension parameter. Similar to the derivation of the 3D straight line rectification, to the pixel p close to the initial curve on an image whose error in the normal direction is dy_n, we have

$$g_T = g\,(y_n + dy_n) \quad (12)$$

$$v_g = \frac{\partial g}{\partial y_N}.\, dy_N - (g_T - g) = g_{y_n} dy_n - l_g \quad (13)$$

and

$$dy_n = \frac{d_y}{\cos\theta} \quad (14)$$

$$\text{or} \quad dy_n = -\frac{d_x}{\sin\theta} \quad (15).$$

g_{y_n} is the gray gradient in the normal direction. From equation (11), the errors, dx and dy, of p in the x and y direction are given by

$$dx = c_0(u) \cdot dx_{i-1} + c_1(u) \cdot dx_i + c_2(u) \cdot dx_{i+1} + c_3(u) \cdot dx_{i+2} \quad (16)$$

$$dy = c_0(u) \cdot dy_{i-1} + c_1(u) \cdot dy_i + c_2(u) \cdot dy_{i+1} + c_3(u) \cdot dy_{i+2} \quad (17)$$

The deformation of a cardinal spline on the stereo images is from the geometric error of the curve in 3D object space. When $|\cos\theta| > |\sin\theta|$, equations (3), (4), (13), (14), (16) and (17) can be used to derive the model rectifying the 3D curve

$$\sum_{k=i-1}^{i+2} c_{x(k)}.\, dX_k + \sum_{k=i-1}^{i+2} c_{y(k)}.\, dY_k + \sum_{k=i-1}^{i+2} c_{y(k)}.\, dZ_k = l_g \quad (18)$$

where the coefficients are

$$C_{x(k)} = g_{yn}\, c_r(u) B_x / \cos\theta \quad (19)$$
$$C_{y(k)} = g_{yn}\, c_r(u) B_y / \cos\theta \quad (20)$$
$$C_{z(k)} = g_{yn}\, c_r(u) B_z / \cos\theta \quad (21)$$

$r = 0, 1, 2, 3$ when $k = i-1, i, i+1, i+2$, respectively. When $|\sin\theta| \le |\cos\theta|$, equations (3), (4), (13), (15), (16) and (17) are used to derive a similar model for the rectification.

Scalable Slope Edge Model

Thus far we have derived the rectification model equations (10) and (18). Assuming the rotation angles (ω, ϕ, x) of the stereo images and the coordinates of the camera centre (X_L, Y_L, Z_L) are known, in the equations the gray value of the template pixel g_T has to be obtained, which is based on a scalable slope edge model and adaptive template generation.

Figure 4 depicts the explicit model of the scalable slope edge profile based on an exponential function, which is profile model along the edge direction. The model is defined as

$$g(s) = h + \frac{k}{1 + \exp(-as)} \tag{22}$$

where h and k are the intensity of the background and the contrast, respectively. a controls the 'scale' of the edge. Larger a results in steeper edge while smaller a means that the edge occurs in a less steep edge. As shown in Figure 3, the strict geometric position of the edge locates at the 'zero crossing' position which is defined by the second derivative of the profile model or the maximum of the first derivative. As discussed in [*Elder and Zucker* 1998], the standard approach to edge detection is based on a model of edges as large-step changes in intensity. This approach fails to reliably detect and localize edges in natural images where blur scale and contrast can vary over a broad range. On remotely sensed images we can find a wide range of edge blur scales and the blur effect of the optical imaging system tends to be the Gaussian blur. This explicit model can be seen as an approximation to the Gaussian edge model employed by other applications [*Elder and Zucker* 1998; *Shan and Boon* 2000; *Ye et al.* 2005].

Adaptive Generation of Edge Template

In order to integrate the edge model into the derived models of linear feature rectification, we have to generate automatically the edge template for each observation pixel. In other words, each edge template is generated locally and adaptively by estimating the three parameters indicating the intensity, contrast and scale of the edge on the pixel.

- Step 1. Generate the default edge templates $g_T(s)$ based on fixed h and k and varying a. For example, setting $h = 50$, $k = 50$ and $a = 2.0$, 1.0, 0.5, we get three edge-profile templates corresponding to three different edge scales.
- Step 2. In the normal direction of the current point of the linear feature, slide a chosen template and do cross correlation with the corresponding image window whose size is the same as the template. Find the maximal correlation coefficient ρ_m and record the corresponding α, let $\alpha_0 = \alpha$. If $\rho_m < \rho_T$ (e.g., $\rho_T = 0.80$), we discard the observation because of the low possibility for the existence of an edge. Note that we need to reverse the template $g_T(s)$ to $g_T(-s)$ and do cross correlation again to find ρ_m which is correspondent to the edge in the opposite edge direction. From the image window g_m correspondent to, ρ_m we can estimate the initial intensity h_0 by averaging the low gray value pixels and the initial contrast k_0 by subtracting h_0 from the average high gray value.
- Step 3. As shown in Figure 5, we use the pixels in the image window g_m for estimating the accurate edge model in the local position. According to the edge model (8), the error between the pixel g_m and the corresponding value in the template g_T is

$$\Delta g = g_T - g_m = \Delta h + \frac{\Delta k}{1 + \exp(-\alpha_0 s)} + \frac{k_0 \alpha_0 \exp(-\alpha_0 s)}{1 + \exp(-\alpha_0 s))^2} \Delta\alpha \tag{23}$$

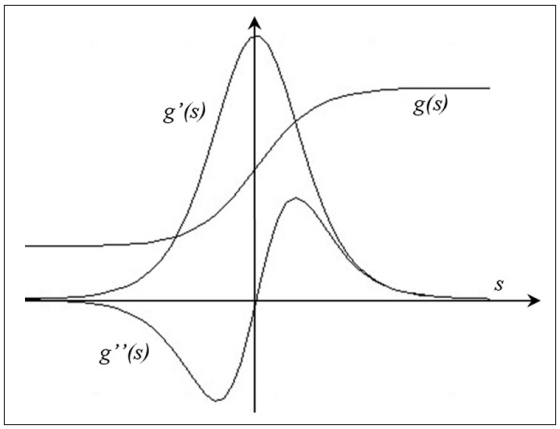


Figure 4: Scalable slope edge model.

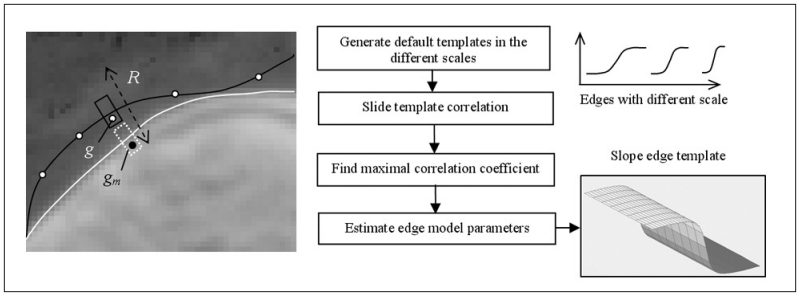


Figure 5: Adaptive generation of edge template.

This is the adjustment model for rectifying the initial parameters (h_0, k_0, a_0) to the least-squares error-based values. The rectified edge profile model forms the template by the updated h, k, and a. The adaptively generated template thus fits to the local intensity, contrast and the scale of the edge.

Experiments on 2D and 3D Using a Commercial Digital Photogrammetric System

The proposed algorithm was tested with 20-cm resolution black and white stereo aerial images. The iterative adjustment procedure is terminated when the maximal position error of the end points of the straight line or the control points of the cardinal spline is less than 0.02 pixel. To adaptively generate the edge templates, the sliding correlation mentioned before has to be done. The search range of the sliding correlation limits the 'pull-in' range of the line rectification. In our implementation we set

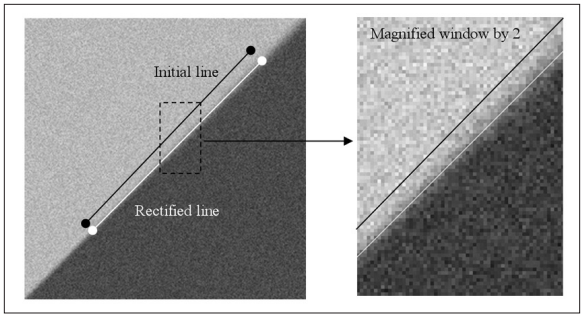


Figure 6: A test image for straight line rectification.

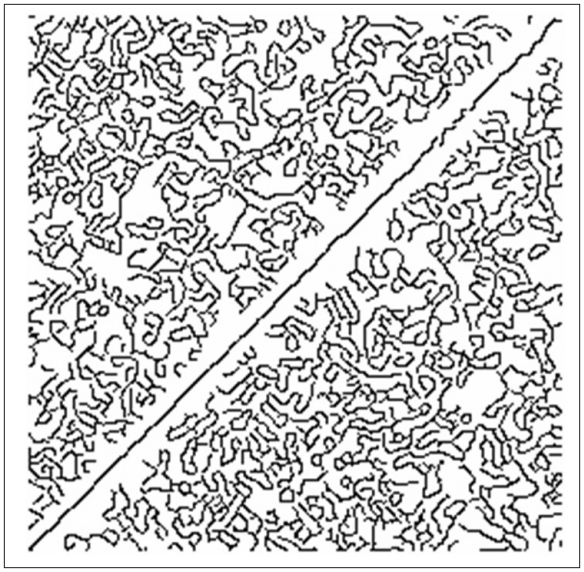


Figure 7: Edge detection to the test image by the Canny operator.

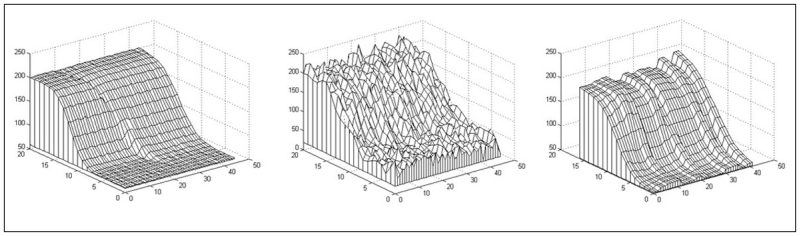
the search range R = 17 pixels in the two opposite directions. We generate the default templates by setting h = 50, k = 30 and a = 3.0, 1.0, 0.6, and 0.4; so the four different scale templates serve as the initial templates. Our edge model is essentially a profile model along the edge direction. Here we choose a W by H size for the generated templates. W = 3 is the width of the template and H is defined by the edge scale. In the implementation, we use a fixed scale of 7 pixels to generate a template, so H = 2x7+1 = 15, and the edge position is centred in the template. The 3 by 15 size template is better than the one-dimensional template in obtaining more reliable results in cross correlation and estimation of the accurate edge model due to the abundant observations. For the cardinal spline extraction, we use the tension parameter t = 0.5.

The proposed algorithm was compared with the edge point fitting method by using a synthetic image. The edge point fitting method is based on minimizing squared errors of the distance from the detected edge points to the line. It is reasonable that we evaluate the accuracy using a single image, because the accuracy of 3D rectification depends on the accuracy of the rectification on each single image and the resolution (scale) of the stereo aerial images. For straight line extraction, we created an original image with the dimension of 256 by 256. A diagonal edge is created by setting pixels at the opposite side as low and high grey value 0 and 255, while setting the grey values on the diagonal line as 128. Apparently the accurate straight line is $y = x$, as shown in Figure 6. We change the contrast of the edge and blur the edge using Gaussian filtering based on various deviations ranging from 0.8 to 4.0 (along the edge and from low-left to upper-right). The additive Gaussian noise is also included in the image, with a mean of zero and a range of standard deviations related to noise ratios (NR) 10% defined by *Ye et al.* [2005]. Fifty initial lines are randomly created around the accurate line position. For each line we employ the two methods to rectify it. For the edge point fitting method, the Canny operator [*Canny* 1986] is used to detect all edge points (see Figure 7). Figure 8 shows an example of the adaptive edge template generation when the image is noisy and the edge scale varies at different locations. The errors of the end points of the 50 inputted initial lines are shown in Figure 9a. For each rectification, the mean error $\bar{d}$ (per pixel) between the rectified line and the real line $y = x$ are computed. For each pixel on the line, the error can be obtained by computing the distance from the point to the real line. Figure 9b shows the comparision of $\bar{d}$ using the two methods. By the proposed method, the

average $\bar{d}$ can reach 0.069 pixel compared to 0.34 pixel error by the fitting method. Simply fitting the detected edge pixels could lead to inaccurate results due to the unreliable edge detection, especially on noisy images and varying slope edge scales. In the proposed method, the edge detection is carried out internally based on template matching by cross correlation, in which the edge scale is modelled. Based on least-squares errors, the edge templates are generated locally and adaptively to fit the edge contrast and scale. The proposed method produces higher accuracy from both single and stereo images, and logically higher accuracy can be expected in 3D rectification.

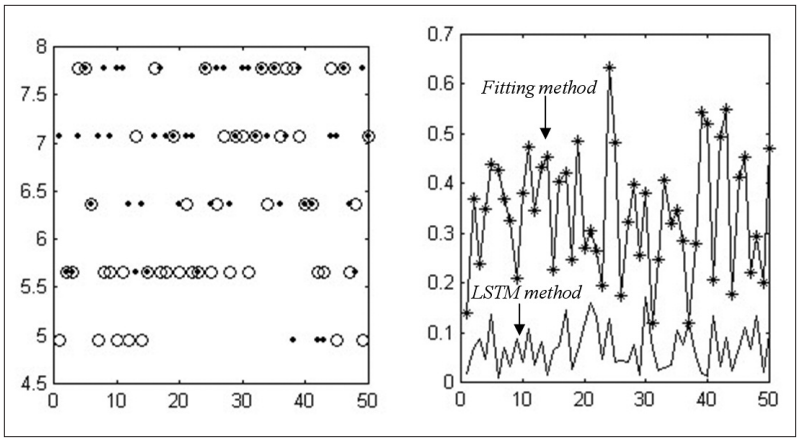
The experiments on 3D line extraction were carried out using a commercial digital photogrammetric system, VirtuoZo, developed by Supresoft Inc, Wuhan, China. The proposed algorithms were implemented based on the derived equations (10) and (18). In order to ease the 3D feature collection from stereo aerial imagery, the system first does relative and absolute orientations (which uses GCPs) to generate an epipolar image pair, from which one only needs to adjust parallax in the direction of image row to get the elevation of a point. A user can input the initial linear feature roughly by clicking a few 'key' points or drag an initial known wireframe to the approximate position on the epipolar image pair. The inaccurate linear feature or the wireframe can be automatically rectified to its accurate position. Note that the geometry of an epipolar image pair and 3D object space is slightly different with the ones expressed in equations (1) and (2); we revised and simplified the final versions of equations (10) and (18) accordingly. In general, equations (10) and (18) are the basis of developing more complex algorithms using LSTM for 3D line extraction.

Figure 10 shows an example of 3D spline rectification from an epipolar image pair. The inputted approximate polygon defines the initial 3D cardinal spline by using the vertices of the polygon as the control points of the spline. It is rectified to the accurate position regardless of the casting of shadows on the edges. The accurate 3D spline can be directly collected, as shown in the zoomed image windows and 3D curve shown in Figure 10a and 10b, respectively.



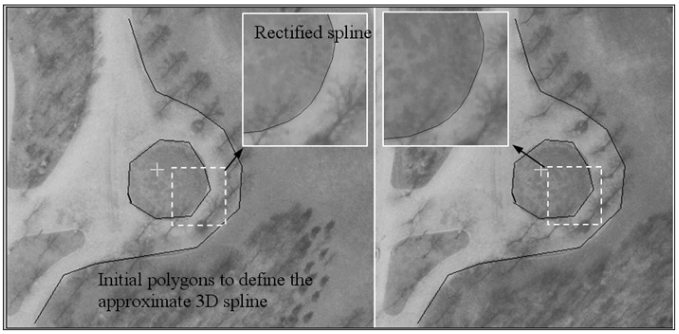
(a) Gaussian blurred slope edge. **(b) 10% NR Gaussian noise added to (a).** **(c) Adaptively generated edge template from (b).**

Figure 8: An example of edge template generation. In (a), (b) and (c) the axis of the horizontal plane is the image row and column, respectively, and the vertical axis is the gray value of the image pixel.

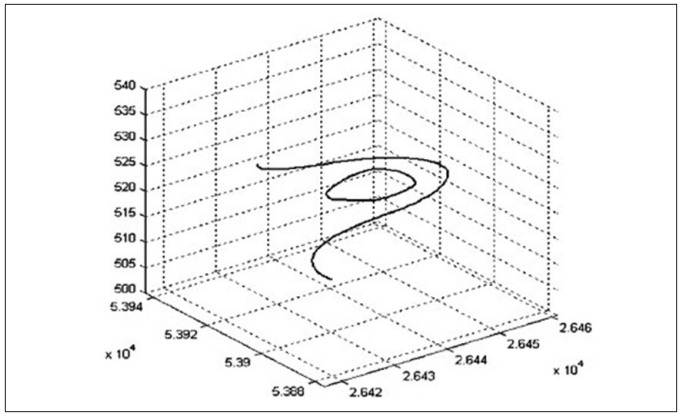


(a) Initial errors of the two end points. **(b) Mean errors of the points on the llines rectified by LSTM and edge points fitting.**

Figure 9: Fifty tests for comparing LSTM method and edge points fitting. The horizontal axis is the times of generating random initial lines, while the vertical axis is the position error in pixel unit.



(a) Inputted approximate polygon and the rectified spline on the images.



(b) Rectified curve in the 3D object space (ground coordinate system).

Figure 10: 3D spline rectification from aerial stereo images.

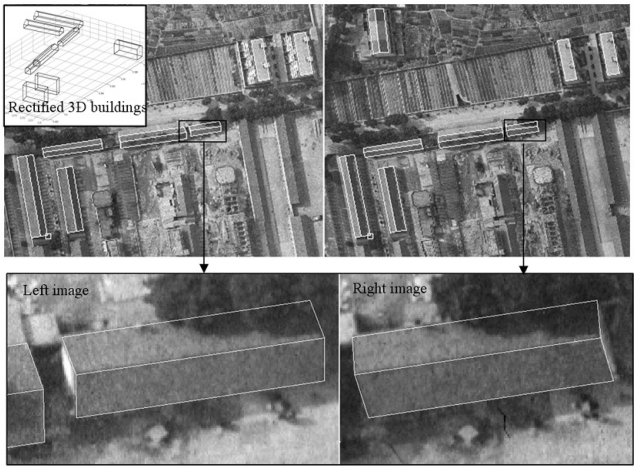


Figure 11: Result of building extraction from stereo images.

A semiautomatic building extraction system is developed by extending the proposed straight line rectification algorithm to extract the edges of buildings. In the system, the initial known wireframe model of the building is input by the user with his visual recognition; the approximate model is then rectified based on the proposed 3D straight line rectification algorithm. How to use minimal input (mouse 'clicking') to obtain initial 3D roof models is another topic we are addressing, using model-driven perceptual grouping explained in a separate paper. Figure 11 illustrates the result of building extraction. The result indicates that the building edges are accurately positioned on the stereo images, overcoming the varying edge blur scales and random noises. In this paper, only visual inspection on the extraction results was used to judge the performance of 3D building extraction, in part because the performance of the proposed semi-automated object extraction method is highly dependant on the user's measurement skill.

Conclusions

Based on 2D analysis of edge models, a new semi-automated 3D feature extraction algorithm has been presented in this paper. The 3D lines are extracted by using template matching based on least-squares correlation adjustment. In order to improve the edge point fitting method, a rigorous edge model with varying scales was utilized, which leads to the higher accuracy of rectification. Examples of 3D line and building extraction are given using a pair of high-resolution aerial stereo images. The results indicate that the derived computational models for 3D rectification of straight lines and splines are feasible and effective. As the imaging geometry of optical images is essentially identical to the image formation of aerial images, the proposed method is capable of dealing with related computer vision problems in which high-accuracy feature extraction is important.

Further research should be carried out in several directions:

1. Speeding up the convergence of the least-squares adjustment. The searching range of the edge matching by cross correlation ('pull-in' range of the adjustment process), formation of the edge templates and the number of the observations are critical elements for convergence. Assigning weights to the observations, using multi-scale method (coarse-to-fine) to do edge matching, simplifying the edge template generation, and reducing the number of the pixels used for constructing the adjustment model—for example, just using one pixel to be an observation in a template could be helpful for faster convergence. We need to find a trade-off of the performance in accuracy and speed.
2. Taking into account more local feature models. There are many more features than slope edges on images. Modeled novel features can be immediately integrated into the proposed model to 3D feature rectification.
3. In the proposed method, we assume that the geometric error of the initial linear feature is from the point error of the line in the normal direction. This may logically lead to the failure of rectification when the initial linear feature is far from shifted deformation. Novel strategies are worth developing to handle these scenarios. Also, error propagations throughout the process, for instance, errors introduced by orientation, were not fully considered in the model. More sophisticated models need to be developed to address all the sources of geometric errors.
4. We have evaluated the accuracy of our algorithms in 2D but not in 3D. Besides visual inspection, quantitative accuracy and precision assessment of 3D line extraction should be developed to evaluate the method.

References

Canny, J. 1986. A computational approach to edge detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 8: 679-698.

Elder, J.H. and S.W. Zucker. 1998. Local scale control for edge detection and blur estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20, 699-716.

Grün, A. and P. Agouris. 1994. Linear feature extraction by least squares template matching constrained by

internal shape forces. *International Archives of Photogrammetry & Remote Sensing,* 30(Part3/1): 316-323.

Grün, A. and H. Li. 1995. Road extraction from aerial and satellite images by dynamic programming, *ISPRS Journal of Photogrammetry and Remote Sensing*, 50 (4): 11-20.

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

Hu, X.Y., Z.X. Zhang and C.V. Tao. 2004. A robust method for semi-automatic extraction of road centerlines using a piecewise parabolic model and least squares template matching. *Photogrammetric Engineering & Remote Sensing*, 70 (12): 1393-1398.

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

Lowe, F. G. 1991. Fitting parameterized three-dimensional models to images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13: 441-450.

Shan, Y. and G.W. Boon. 2000. Sub-pixel location of edges with non-uniform blurring: a finite closed-form approach. *Image and Vision Computing*, 18 (13): 1015-1023.

Steger, C. 1998. Unbiased Extraction of Curvilinear Structures from 2D and 3D Images. Ph.D.Dissertation, Technical University of Munich, Germany, 192 pp.

Steger, C. 2000. Subpixel-precise extraction of lines and edges. International Archives of Photogrammetry and Remote Sensing, 33, Amsterdam, The Netherlands, pp. 141-156.

Trinder, J. and H. Li. 1995. Semi-automatic feature extraction by snakes, In: Grün A., O. Kübler, and P. Agouris (eds.) *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, Birkhäuser Verlag, Basel, pp. 95-104.

Tseng, Y.H. and S. Wang. 2003. Semi-automated building extraction based on CSG model-image fitting, *Photogrammetric Engineering and Remote Sensing*, 69(2): 171-180.

Vosselman, G. and H. Veldhuis. 1999. Mapping by dragging and fitting of wire-frame models, *Photogrammetric Engineering and Remote Sensing*, 65: 769-776.

Wolf, P.R. and B.A. Dewitt. 2000. *Elements of Photogrammetry with Applications in GIS*, 3rd edition, McGraw-Hill, Toronto, 608 pp.

Ye, J., G.K. Fu and U.P. Poudel. 2005. High-accuracy edge detection with blurred edge model. *Image and Vision Computing,* 23 (5): 453-467.

MS rec'd 07/11/14
Revised MS rec'd 08/05/07

Authors

Xiangyun Hu received his Ph.D. degree in photogrammetry and remote sensing from Wuhan University, China in 2002. He held a post-doctoral fellow position at the GeoICT Lab of York University, Toronto, Canada, before joining ERDAS Inc. U.S. as a senior software engineer. He developed a software tool named SmartDigitizer which is based on a series of intelligent algorithms used for fast feature extraction from remotely sensed imagery. He has published many research papers in refereed journals, books, and proceedings. His current research interests are mainly advanced image processing of geospatial imagery and computer vision.

Zhxun Zhang is a professor in photogrammetry and remote sensing at the School of Remote Sensing Information Engineering, Wuhan University, China. He received his B.Eng. degree in photogrammetry from the Wuhan Institute of Surveying and Mapping in 1960. He developed the well-known digital photogrammetric system – VirtuoZo, which has been widely used in the photogrammetric mapping industry worldwide. He was the founding Director of the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing (IESMRS) and former Vice-President of Wuhan Technical University of Surveying and Mapping (WTUSM). He was elected as an academician of the Chinese Academy of Engineering in 2003. He is author of more than 100 research papers. His research interests are digital photogrammetry and computer vision.

Jonathan Li holds a Ph.D. degree in remote sensing and GIS from the University of Cape Town, South Africa. He has been an associate professor in geomatics with the Department of Geography and Environmental Management at the University of Waterloo since January 2007. Before taking this position, he was an associate professor in geomatics engineering at Ryerson University in Toronto. Dr. Li is a Professional Engineer licensed with PEO and an Ontario Land Surveyor licensed with AOLS. He has served on the CIG Photogrammetry and National Committee for ISPRS since 2002. He is co-chair of the ISPRS WG IV/8 (2004-2008) and ICA Commission on Mapping from Satellite Imagery (2008-2011). He received the 2008 ESRI Award for Best Scientific Paper in GIS, and the 2007 MAD Award for Best Technical Paper in Orthophotography, Photogrammetry and Surveying. He has co-edited four books and co-authored about 100 research papers and reports dealing with digital mapping from image and range data, terrain analysis, digital image processing algorithms for automated object extraction and change detection, remote sensing and Internet GIS applications for disaster management. ❑