# ROAD UPDATING FROM HIGH RESOLUTION AERIAL IMAGERY USING ROAD INTERSECTION MODEL

Go Koutaki [a], Keiichi Uchimura [a], Zhencheng Hu [b]

[a] Graduate School of Science and Technology, Kumamoto University, 2-39-1 Kurokami, Kumamoto, 860-8555 Japan - koutaki@navi.cs.kumamoto-u.ac.jp, uchimura@cs.kumamoto-u.ac.jp
[b] Faculty of Engineering,, Kumamoto University,2-39-1 Kurokami, Kumamoto, 860-8555 Japan- hu@cs.kumamoto-u.ac.jp

**ISPRS WG V/6**

**KEY WORDS:** Road extraction, Intersection detection, Snakes

**ABSTRACT:**

We propose an intersection model and strategy for road extraction from aerial image. Proposed approach is able to detect typical intersections such as cross road, T-junction and Y-junction based on the model matching to the image features. The road network is constructed by connecting the detected intersections. The connecting hypothesis is generated and validated using the road tracking method and the road shape including the width is refined using the ribbon snakes. We show the feasibility of our approach by presenting results at the suburban area, and evaluating them comparing to the existing digital road map.

## 1. INTRODUCTION

It is important to shorten the cycles of updating and acquiring new map information. A digital topographic database is an essential part of Geographic Information System (GIS). The building outline or road arc data is demanded in the field of Intelligent Transport System (ITS) such as car navigation, forwarding agencies, city planning or etc. Major existing digital map is digitized from the topographical 1:25,000 paper-based map. It may miss road arc and contain geometric spatial errors because of the scanning error or distortion of paper map. Furthermore, the map may be old since the secular change of urban area is vigorous. Thus post editing with a human operator will be required. Since the manual extraction from the images is time consuming and tedious, automation for updating map information should be employed. Aerial or satellite imagery is an important data source for acquiring topographic objects with high accuracy.

An early road extraction approach is focused in low-resolution aerial images. The road detector considering local and global criteria is proposed (Fischler et al 1980). Road tracing step exploits local criteria calculated by low level processing. The method of line extraction based on a differential geometry is presented (Steger 1996). For each pixel in the image convoluted with the Gaussian kernel, the image profile along the principal direction is examined. Line points that the first and second derivation of the profile have respectively a vanishing and minimum are detected and connected. As approaches for full-automatic road extraction, the multi-resolution approach (Mayer et al 1997) is proposed. In coarse scale in the image, the lines are extracted using the line extraction algorithm of Steger. In fine scale, salient roads with a low variance of width are verified by ribbon snakes (Fua 1990) initialized by results of the line detection in the coarse scale. Non-salient roads that include a shadow, car, or have no road side, are bridged with so called “Ziplock Snakes” (Neuen 1996). A graph structure for the road network (Wiedemann 1997) is developed. Road network is constructed by calculation of the optimal path between road segments with high evidence. Weight of the graph is evaluated with fuzzy value that binds some criterions, i.e. length of segments or gap, straightness and curvature.

Above works have concentrated on open rural areas. Multi-resolution approach mainly depends on the result of line detection in low-resolution. Most roads are distinct from background objects (mostly field and vegetation) and have road sides clearly presented. Therefore, because the line detector or ribbon snakes extract a lot of salient roads very successfully, the construction step of road network works well even if some gaps derived from shadows or tree exist. However, road sides in suburban areas are absence because of house, shadow or bush. Additionally, some house roof that have parallel edges will be extracted as road sides and they will cause false connection for the road network generation.

The road grid model (Price 1999) have been attempted to detect roads in suburban and urban areas. Road network is modeled as a regular street grid patterns with topology and individual roads are approximated with constant width and length. The initial grid is given by a human operator and propagated across the scene. However the grid description only deal with intersection with mostly right angles and regular intervals. It is difficult to extract curve roads.

We propose to detect various intersections aligned with non-regular intervals such as cross-road, T-junction and Y-junction (three-forked road) automatically. They are extracted by model matching method like a template matching, and redundant results are reduced using overlapped resolution. Road network is constructed by connecting the detected intersections. A hypothesis of the connecting is generated and validated using the road tracking method and the road shape including the width is refined using the ribbon snakes. We show the feasibility of our approach by presenting results at the suburban area, and evaluating them comparing to the existing road map.

## 2. DATA SOURCE AND IMAGING SYSTEM

Input aerial images are provided the Digital Earth Technology. The special resolution is 50cm/pixel and they have three channels of RGB. Our approach is summarized in Fig.1. The developed system consists of three parts which are the feature extraction, the intersection detection, and the road network construction. They contain some image analysis tools. The basic idea is connecting between each intersection for road network construction. Because it is rare to isolate road, most roads will be extracted. Contrast to various forms of representation for roads, i.e. curve road, klothoid curve, and a road with changing width, the intersections has similar simple shape. For instance, the roads entering intersection are mostly straight road because of traffic safety or regulation. We consider the geometric intersection model, such as three or four straight roads entering a certain point. The road intersections fit the models are extracted. Certain features are extracted from images, and intersections and roads are extracted using the features. Those processes depend on "Road seed binary image". In the following sections, we illustrate each step in detail.

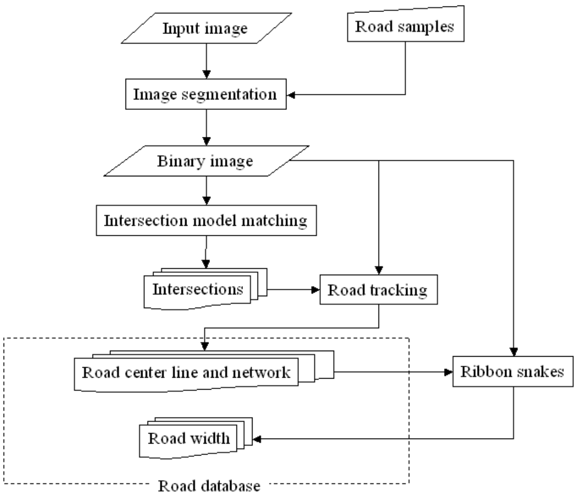


Figure.1 Imaging system for road extraction

## 3. ROAD SEED EXTRACTION

Road seed pixels are segmented by thresholding in color space. A human operator gives a road samples and distribution (mean vector and a covariance matrix) of the samples is exploited for the segmentation. In order to separate road pixels from background efficiently, we transformed the RGB color spaces into three different color spaces, i.e. luminous, greenness in RGB space as $(G\text{-}R)/(G+R)$, and saturation from HIS color space. The greenness allows to enhancing the vegetations, and the saturation is able to emphasize the shadows and house roof with vivid color (Zhang et al 2000). The road pixels are classified in the feature vector space. Fig.2 shows the plotted feature vector of input image. An ellipse in the figures is the estimated normal distribution from road samples that the operator gives. Fig.3 shows the result of the classification for input image. We call the results "road seed".

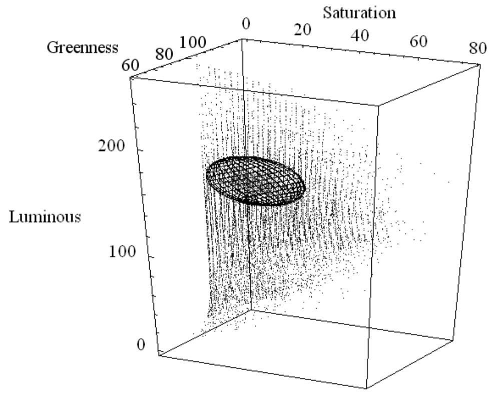


Figure.2 Distribution in the transformed color space

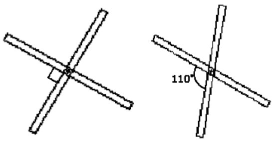
Figure.3 Result of image segmentation
Left: input image
Right: road seed

## 4. INTERSECTION MATCHING

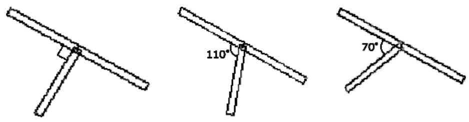
A road seed is binary image that white pixel denotes high probably road like object. Since an extraction error arises from some pixels on building roofs or soil have similar spectral response to roads, general morphological operator, for example the combination with closing, thinning and 8-neighbour pattern matching, will not work well because of very sensitive for noise. Therefore stronger constraint and knowledge about intersection are required. We consider intersection to following three types:

1. the Crossroads represent the intersection of two road portions.
2. the Three-forked road has three road segment. Each branch has different direction.
3. the T-Intersection consists of one straight road and connected branch.

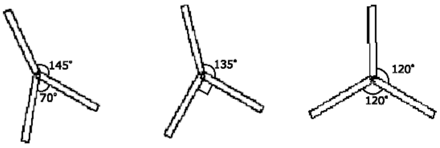
We provide intersection models as templates shown in Fig.4, they consist of two or three elongate rectangles with different or same width each other. Each branch has 4m, 7m, or 13m width corresponding to one lane, two lanes, and three lanes road respectively with constant length $L$=30m. Fig.5 shows models of combination with the widths. We represent the all models as a set $M$. For $M \in M$, $M^{\theta}_{(x,y)}$ denotes that center position of intersection is $(x,y)$ and direction of 1st branch is $\theta(0 \leq \theta < 2\pi)$.




(a) Cross roads




(b) T-junctions




(c) Three-forked roads

Figure.4 Type of intersections

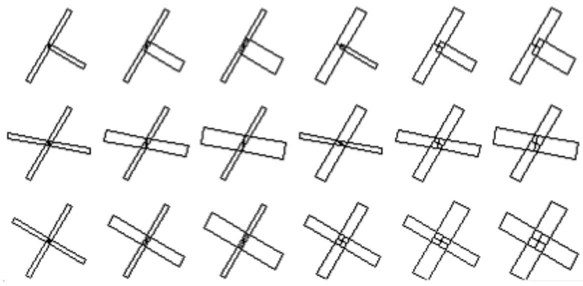


Figure.5 Combination of various widths

Consider matching above models to road seed and calculating matching value between the models and road seed. The model is rotated and positioned over the binary image. The matching measure is defined following,

$$M^{\theta}_{(x,y)} \equiv \begin{cases} \mu(S)-\mu(B), \text{ if } \min\limits_{n=1,2,\dots,N} \mu(S_n) > k_1 \\ 0, \text{ otherwise} \end{cases} \tag{1}$$

Here, $S$ and $B$ is a region of inner and outer model. $S_n$ denotes region inside $n$th branch of intersection and $S=S_1 \cup S_2 \cup \dots \cup S_n$. The constraint of equation (1) requires that whole branch should have values more than threshold $k_1$. $\mu(.)$ is the ratio to the areas of the number of the road seed in the region. It is defined as,

$$\mu(R) = \frac{1}{|R|}\sum_{(x,Y)\in R} I(x,y) \tag{2}$$

$I(x,y)=\{1,0\}$ is the binary image of road seed and $|R|$ denotes the number of pixels in region $R$. Background region $B$ is defined subtracted $S$ from circular region $C$ (radius $L$) in shown Fig.6. The matching algorithm is as follows. For each position $(x,y)$ in input image, the maximum matching measure is calculated following,

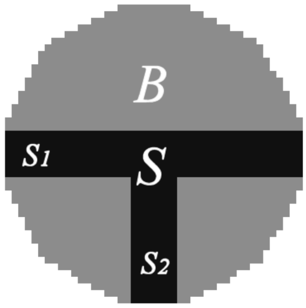
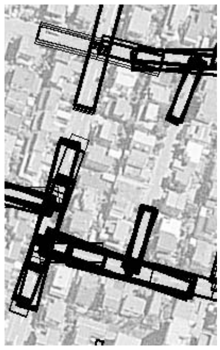



Figure.6 Inner and outer region of intersection model

$$\begin{aligned} D(x,y) &= \max_{M\in M, 0\le\theta<2\pi} D(M^{\theta}_{(x,y)}), \\ M^{\max}_{(x,y)} &= \arg\max_{M\in M, 0\le\theta<2\pi} D(M^{\theta}_{(x,y)}) \end{aligned} \tag{3}$$

If $D(x,y) > k_2$, $M^{\max}_{(x,y)}$ is extracted as intersection and put in the intersection cue $P$. It is available to use "Coarse-To-Fine" matching Firstly, The algorithm tests the matching by not whole pixels but the interval $q_1$. If $D(x,y) > k_3$,in the pixels in the neighborhood of $(x,y)$, the matching is performed by the interval $q_2$. We set $q_1$ to the pixel value corresponding to 2m, and $q_2=1$. Typically, a number of pixels around the center of intersection will have a high matching value, thus many intersections will be detected. Thus $P$ contains redundant intersection. Overlap resolution is needed to remove redundant results. First, we group adjoining intersections that have similar shape and direction. Then, intersection with the highest evaluation value in the group is represented. Overlap resolution of intersections is shown as a Fig.7.



(a) Results of matching

(b) After overlap resolution

Figure.7 Results for overlap resolution

## 5. ROAD NETWORK CONSTRUCTION

### 5.1 Extraction of road center line

Road network is constructed by connecting branches of each intersection. A connection hypothesis is performed according to directions and distance between two branches. Road tracking method (Geman et al 1996 and Zhao et al 2002) is available for the hypothesis. A structure of road curve-linear is modeled as ternary tree (Fig.8). We set $H$=30m and $\alpha$=5°. A starting point and direction of the tracking are given by center point and direction at each branch of the intersection. A part of path are evaluated by matching an elongate rectangle template to road seed as follow equation,

$$E(a) = \mu(A_{in}) - \mu(A_{out}), a \in A \tag{3}$$

Where $A$ is a set edge of the tree. $A_{\mathrm{in}}$ and $A_{\mathrm{out}}$ are inside and outside regions respectively around the edge as shown Fig.9. The size of this template is same as the branch rectangle of the intersection. A road line is extended from starting point in an iterative way by calculating maximum cost. Using tracking approach, road network is generated according to the next three rules.

[R1] if road tracker from a branch can reach the other intersection, the result is inserted as road arc in network.
[R2] While tracking, if tracker has met with previous extracted road arc.
[R3] even if tracker does not fulfill the conditions above, it is possible to insert as road arc in the case of 30m or more of tracking.

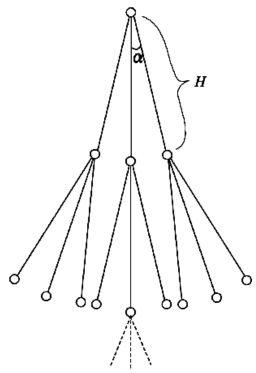



Figure.8 Road structure for tracking

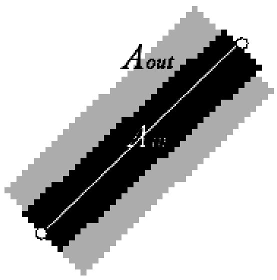



Figure.9 Line template for tracking

### 5.2 Extraction of road width

To refine the road shape containing the width, we implement ribbon snakes (Fua 1990) method that is kind of active contour model (Kass et al 1988). The ribbon snake is initialized using road arc derived in road network construction step. We represent ribbon snakes as the parametric curves to following equation.

$$v(s)=(x(s),y(s),w(s)) \tag{4}$$

Where $s$ is proportional to arc length, $x(s)$ and $y(s)$ is center coordinates of ribbon, and $w(s)$ is the curve's width. The total energy of the ribbon snakes can be defined as,

$$E_{total} = \int_s E_{\mathrm{int}}(v(s)) + E_{\mathrm{img}}(v(s))ds, \tag{5}$$

$$E_{\mathrm{int}}(v(s)) = \alpha\left|\frac{\partial}{\partial s}v(s)\right|^2 + \beta\left|\frac{\partial^2}{\partial s^2}v(s)\right|^2. \tag{6}$$

Here, $\alpha$ and $\beta$ is control parameter of ribbon's rigidness and smoothness. Typically, $E_{\mathrm{img}}$ is defined by intensity or gradient of image. However, in order to embed in road seed as image energy, we define total energy as follows,

$$E_{total} = \int_s E_{\mathrm{int}}(v(s))ds + E_{\mathrm{img}}(R_{\mathrm{in}}, R_{\mathrm{out}}) \tag{7}$$

$R_{\mathrm{in}}$ and $R_{\mathrm{out}}$ are the regions corresponding to inner and outer of the ribbon. In order to perform the numerical optimization of the energy, equation (7) is discretized. $E_{\mathrm{int}}$ is calculated in finite difference. $R_{\mathrm{in}}$ and $R_{\mathrm{out}}$ is approximated as sequential rectangles, $\{R_{\mathrm{in}}^1, R_{\mathrm{in}}^2, \ldots, R_{\mathrm{in}}^N\}$ and $\{R_{\mathrm{out}}^1, R_{\mathrm{out}}^2, \ldots, R_{\mathrm{out}}^N\}$ as shown in Fig.10. The discretized version of $E_{\mathrm{img}}$ is defined as follows,

$$E_{\mathrm{img}}(R_{\mathrm{in}}, R_{\mathrm{out}}) \approx \lambda \sum_{i=1}^{N} \left(\mu(R_{in}^i) - \mu(R_{out}^i)\right) \tag{8}$$

Where $\lambda$ is control paramer of image energy. It's possible to optimize the total energy using the dynamic programming (Amini 1990). Fig.11 shows the example of ribbon snakes based on road seed.

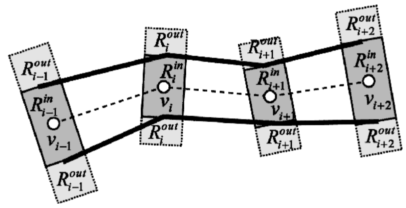


Figure.10 Region of ribbon for image energy

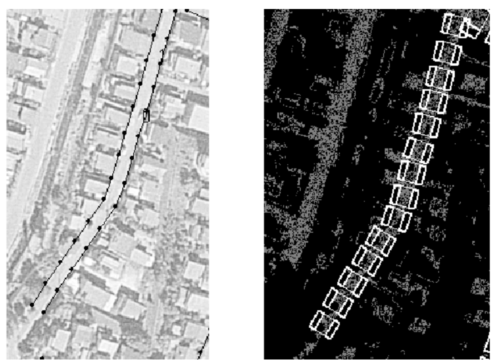


Figure.11 Results for ribbon and aspects in road seed

## 6. RESULTS

Our approach has been tested in one rural area and two suburban areas. The size of these input images is 1000×1000 pixels corresponding to 500×500m. We set the parameters as $k_1$=0.3, $k_2$=0.25, $k_3$=0.125, $q_1$=4, $q_2$=1, and as ribbon's parameters, $\alpha$=1.0, $\beta$=2.0, $\lambda$=3.0 is used. For the reasons of comparison the test was performed with multi-scale snakes (MS) method (Ivan 1997). The results have been evaluated by matching the extracted road axes to Digital Map 2500 that the Geographic Survey Institute Japan published (1/2500 scale). The buffer method (Heipke 1997) is used for the quality measures. Note that we compared in respect to center lines since the reference data did not contain width. Buffer size is used as 15 pixels. Table.1 and Table.2 show quality of results. The results was obtained on a PC with Pentium III processor 2.8 GHz, and 1024 MB of RAM.

In rural area, MS can extract road network correctly since line extraction in coarse scale works well. However in the roads adjoining soil, line and edge extraction fell down because of low contrast. For the same reasons, our approach could not detect the roads of same areas, too. In suburban areas, MS fell down to extract road network considerably since line and parallel edges could not be extracted because of house roof, junction or bush. Further some road segments are extracted unfortunately in building roof. It results in extracting false road connection. On the other hand, our approach could extract intersections correctly and not extract buildings as roads even if in suburban areas. As the results, we could construct road network with high correctness.

In order to validate the proposed approach, we tested in the larger image in suburban area (the size is 2000×2000 pixels). Fig.12 shows the results for the image. Here, black ribbons show extracted road sides and black dot lines mean the roads that our system could not extract. The quality of this image was 68.7%. Fig.13 shows the zoomed part of the upper-right of Fig.12. It is the parts of including several failures especially. These parts are numbered serially and the each part is as follows.

①The crossing with an overpass or underpass

②There exists in the reference data, but the roads do not exit in image.

③The wide roads and intersections with 6 lanes and markers

④The roads in privately owned area

⑤They are only extracted as the pavement.

⑥L-intersections

⑦The intersections adjoined a parking area or house roof

Table.1 Results for MS method

| | Rural | Saburban 1 | Saburban 2 |
|---|---|---|---|
| Completeness [%] | 50.2 | 43.1 | 42.9 |
| Correctness [%] | 98.4 | 57.9 | 69.7 |
| Quality [%] | 49.4 | 25.0 | 29.9 |
| Time [minite] | 20.2 | 40.5 | 46.6 |

Table.2 Results for proposed method

| | Rural | Saburban 1 | Saburban 2 |
|---|---|---|---|
| Completeness [%] | 39.0 | 72.8 | 93.2 |
| Correctness [%] | 90.9 | 95.5 | 95.7 |
| Quality [%] | 35.5 | 70.2 | 89.2 |
| Time [minite] | 14.4 | 15.2 | 16.5 |

## 7. CONCLUSION

We proposed the intersection models for constructing road network from aerial images. The models are classified to three types of crossroads, T-junctions and Y-junctions. Road network is generated by connecting road intersections using the road tracking method. Road width is modified by ribbon snakes method that is based on not the intensity of image but a binary image.

The results show that the proposed method is useful for extracting roads in the residential areas that have typical intersections. However, several roads adjoining the parking areas or the buildings were not extracted since intersection matching was failed in those places. It is difficult to extract these roads because the parking and building have similar color to the road's surface. We should use not only the features of binary image corresponding to the road's surface but also the features of lines and edges of image.

Figure.12 Final results for our approach

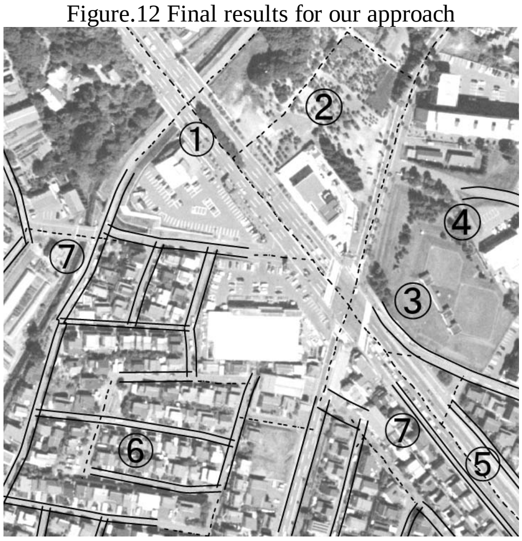



Figure.13 Failure parts

**References:**


Amini, A. 1990. Using Dynamic Programing for Solving Variational Problems in Vision. *PAMI*, 12(9): pp. 885-867

Fischler, M. Tenenbaum, J. and Wolf, H. 1981. Detection of Roads and Linear Structures in Low Resolution Aerial Images Using Multi-Source Knowledge Integration Techniques. *CGIP*, 15(3): pp. 201-223

Fua, P. and Leclerc, Y.G. 1990. Model Driven Edge Detection. *Computer Vision*, 3(1): pp. 45-56

Geman, D. and Jedynak, B. 1996. An Active Testing Model for Tracking Roads in Satellite Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(1): pp. 1-14,

Heipke, C. Mayer, H. and Wiedemann, C. 1998. Empirical Evaluation of Automatically Extracted Road Axes. *IEEE Computer Society Press, Empirical Evaluation Methods in Computer Vision*: pp. 172–187

Kass, M. Witkin, A. and Terzopoulos, D. 1988. Snakes: Active contour models. *Computer Vision*, 1(4): pp. 321-331

Laptev, I. 1997. Road Extraction Based on Snakes and Sophisticated Line Extraction. *Master Thesis, Computational Vision and Acive Perception Lab(CVAP), Royal Institute of Technology in Stockholm, Sweden*

Laptev, I. Mayer, H. Lindeberg, T. Eckstein, W. Steger, C. and Baumgartner, A. 2000. Automatic extraction of roads from aerial images based on scale-space and snakes. *Machine Vision and Applications*, 12(1): pp. 23-31

Neuenschwander, W. 1996. Elastic Deformable Contour and Surface Models for 2-D and 3-D Image Segmentation. *Hartung-Gorre Verlag*, 1996

Price, K. 2000. Urban Street Grid Description and Verification. *In Proc. Of 5$^{th}$ IEEE Workshop on Application of Computer Vision*: pp. 148-154

Steger, C. 1998. An Unbiased Detector of Curvilinear Structures. *PAMI*, 20(2): pp. 113-125

Wiedemann, C. and Hinz, S. 1999. Automatic Extraction and Evaluation of Road Networks from Satellite Imagery. *International Archives of Photogrammetry and Remote Sensing,* 32: pp. 95-100

Zhang, C. Baltsavias, E. and Gruen, A. 2000. Knowledge-based image analysis for 3D road reconstruction. *Proceedings of the 21st Asian Conference on Remote Sensing*, 1: pp. 100-105

Zhao, H. Kumagai, J. Nakagawa, M. and Shibasaki, R. 2002. Semi-Automatic Road Extraction from High-Resolution Satellite Image. *Proceedings of Photogrammetric Computer Vision*: pp. 406-411