

A SEMI AUTOMATIC ROAD EXTRACTION METHOD FOR ALOS SATELLITE IMAGERY

H. Hasegawa

Geographical Survey Institute, 1 Kitasato, Tsukuba, Ibaraki, 3050811, Japan – hase@gsi.go.jp

KEY WORDS: Semi-automation, High resolution Satellite, Edge Recognition, Multi sensor Interpretation, GIS, Cartography

ABSTRACT:

A semi automatic road extraction method has been constructed and evaluated. Our purpose is to obtain 1/25,000 level road data from PRISM image. PRISM is panchromatic, three views, and the 2.5m resolution sensor carried on ALOS satellite that will be launched on late 2004. A centre line-detecting algorithm is employed for feature extraction. It picks up pixels where second derivative of brightness becomes maximal. Then those pixels are linked if both probability of line and angle difference between adjacent pixels satisfy given conditions. Acquired line candidates are classified by its photometric property. We test both automatic grey scale threshold method and traditional unsupervised multi band classification method for this stage. After eliminating false line segments, a line linking method is applied. If all of angle difference, lateral offset, and net gap are less than threshold, the pair is considered as one long line. This process is applied iteratively while a connectable pair remains. The proposed method was tested by a simulated ALOS image data set created from three-line airborne sensor images and an IKONOS data set.

The result shows that 80% of road was extracted before false line elimination while 80% of extracted line was false data. There still remains 70% of false road segments even after the classification method. Both correctness and completeness are unexpectedly poor. As our method is working on a single image and does not use full feature of PRISM yet, a method using three-dimensional property is needed.

1. INTRODUCTION

This research focuses on extracting 1/25,000 level road network from ALOS optical sensor data. Geographical Survey Institute of Japan (GSI) is the national mapping agency in Japan and GSI has responsibility to keep 1/25,000 level map data the latest.

The data acquisition method, however, has mainly depended on manual aerial photo interpretation. As a result, map revision interval is too long (3-5yr in urban and suburban area, 5-10yr in rural area) to follow rapid changes especially in urban area.

Advanced Land Observation Satellite (ALOS) will be launched at the end of 2004. Its primary task is making map in Japan. ALOS is expected to another source for map making. However, as data amount is enormous and data generation from optical images is time consuming and expensive task, a semi automatic road network extraction system is needed.

A great amount of research has been targeted on developing methods to extract features like roads and buildings. As a full automatic feature extraction method is obviously out of reach, many investigations have been focusing on semi automatic feature extraction method.

Baltsavias (2002) briefly reviewed the recent trend of image analysis, feature extraction strategy, and existing commercial systems regarding automated object extraction. The paper pointed out that proper definition of the target objects, choice of adequate input data, use of existing knowledge, and different data analysis method from low level to high level was the main considerations. It argued that object modelling had become popular as level of modelling was highly depends on desired resolution and that combined use of sensor data was more common. It also said use of contextual information such as relationship between neighbours and usage of as much cues as possible (i.e. use multi spectral data for land cover classification and use DTM for shadow analysis) had received much attention.

Baumgartner et al (1999) proposed an automated road extracting system that used hierarchical road model applied to

aerial photo (0.2 – 0.5m resolution). Based on their previous work, they clearly distinguished global contexts (rural, forest and urban) and corresponding local contexts were selected and used for road extraction process. Their approach worked quite well in rural area with less than 0.5m resolution images whereas more improved extraction and grouping methods were needed in urban and forest area.

Price (1999) proposed a road grid extraction method standing on a specific assumption. They assumed that streets crossed regularly at right angles. The result was quite fine and processing speed was very rapid although their assumption may not be applicable to complex urban scene.

Zhao et al (2003) proposed another semi-automatic road extraction method using multi spectral high-resolution satellite images. Firstly “road mask” was created by multi spectral data classification. Chains of edge pixels were tracked based on local edge direction and straight lines were obtained. Then this method use template matching to determine the best direction of line and obtain next road node. A result in urban area was good for major road whereas small road were missed, as road boundaries were unclear due to the objects surrounding roads. In rural area, both major and minor roads were properly extracted by indicating adequate control points.

Treash and Amaratunga (2000) proposed an automatic road detection system using high-resolution greyscale images. This system consists of three parts, edge detection, edge thinning, and edge pairing. The result is good in rather simple image whereas in highly occluded road and dense residence area it shows poor performance. The authors suggest some additional information based on texture of grey value is needed.

Park and Kim (2001) proposed a template based road extraction system. A user needs to input an initial seed point to extract a road. Then orientation of the road seed point is calculated automatically. They conclude that this template matching method have more flexibility than the methods using snake. They pointed out the method may not work on the road cast by shadow.

Hu and Tao (2002) proposed template matching based main road extraction method for high-resolution satellite image. A two direction (horizontal and vertical) template matching is applied to the reduced resolution image. In line grouping process, they classified line segments and the segment connectivity matrix is constructed. Then collinear chains are extracted. Finally, precise position of road centre line is extracted on the original image by least square template matching. They applied their method on a IKONOS image. Main road is correctly extracted in the open rural area while the result in the complex building site is not so good because of the complicated road network.

According to Vosselman & Knecht (1995), road characteristics can be classified in five groups; Photometric, Geometric, topological, functional, and contextual characteristics. As ground resolution of PRISM data is 2.5m and small marks on a road are not visible, precise road model are not adequate for this study. Instead of a precise model, simple line detection methods, which use photometric and geometric characteristics, are employed for road extraction in this study. Geometric, photometric and contextual characteristics are used for grouping line segments and building road network topology.

In this paper, we propose a semi-automatic road extraction method based on three stages; line feature extraction stage, line segment classification stage, and line segment grouping stage. A centre line-detecting algorithm proposed by Steger (1998) is employed for feature extraction. It picks up pixels where second derivative of brightness becomes maximal. Those pixels are linked and chains of pixels are created if both probability of line and angle difference between adjacent pixels satisfy given conditions. In line segment classification stage, acquired line candidates are classified by its photometric property. We test both automatic grey scale threshold method and traditional unsupervised multi band classification method for this stage. After eliminating false line segments, a line linking method is applied on the line segment grouping stage. If all of angle difference, lateral offset, and net gap are less than threshold, the pair is considered as one long line. This process is applied iteratively while a connectable pair remains.

In the following section, methods for each stage are described. Then we present extraction results for each stage. Finally, we discuss the result and give direction for further work.

2. METHODS

2.1 Data used

A simulated ALOS PRISM data set has been used for this study. It has been offered by NASDA/EORC for limited use. Data was acquired with an airborne three-line CCD sensor to simulate along-track PRISM sensors. Original data (25cm resolution) was thinned down to fit the targeting resolution (2.5m) by averaging surrounding 10x10 pixels. Ancillary data such as sensor position and attitude corresponding each image were also provided, though image data is only used for this study.

As the image size is too large to manipulate, some target area are selected. Figure 1 (a) and (b) are used for evaluation. Figure 1 (a) is 329 by 642 pixels and dense residential area. It contains some residential quarters and some single lane or dual lane roads divide them. Dual lane roads separate residential area and its adjacent area such as forest and agricultural area. Figure 1 (b) is 300 by 600 pixels. Most part is paddy field and dry field. A major road goes through the northward. Some habitations can be seen at the east area and minor roads go through them. All roads are clear to human eyes. Ground truth data is created from

1:25,000 vector map data and the simulated images by manual compilation.

As the data set only include the simulated PRISM images but not contain simulated AVNIR-2 image, both PRISM and AVNIR-2 simulation data were created from IKONOS images to investigate effectiveness of multi band classification. Standard geometrically corrected IKONOS 4band image in Farnborough, Hampshire, the United Kingdom and IKONOS panchromatic image in the same area with the same process level were used. Both images are firstly reduced in resolution to fit the ALOS resolution. The resolution merge tool of ERDAS was used to create the pan-sharpen image. The merging method is the principal component method and the re-sampling technique is the cubic convolution method. Its image size is 400 by 400 pixels. Main roads are clearly recognized as grey elongated area.

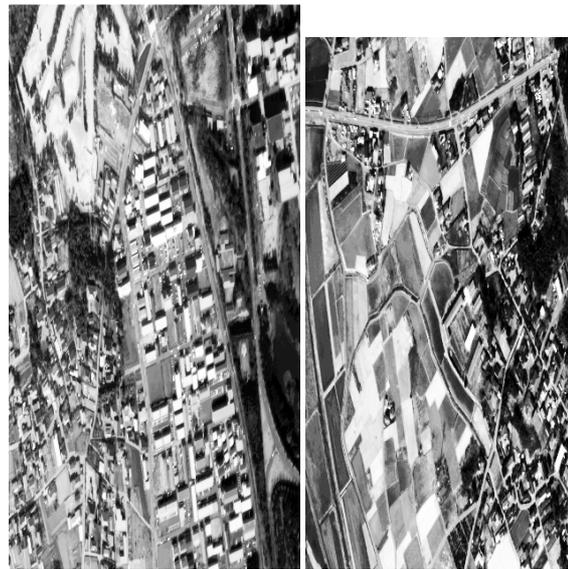


Figure 1. Target Images for dense residence area (a) and for rural area (b). The image size is 329 by 642 pixels (a) and 300 by 600 pixels (b). Histogram equalisation was applied for both images.

2.2 Line Feature Extraction Stage

Many studies employ edge detector such as Canny, SUSAN for feature extraction method. A ridge extraction method proposed by Steger (1998) is used in this study. It follows centre of bright (or dark) blob where the second derivatives of profile crosses zero. The method gives centre (not edge) of road, needs no excess parallel edge process, and has sub-pixel accuracy. It consists of two stages, line point finding stage and line linking stage. It assumes that roads are almost homogeneous and have clear contrast to their adjacent areas. However, as real images have some noises that give us false ridge, it used Gaussian smoothing kernel for convolution for noise reduction in line point finding stage. Characteristic of this method is, smoothing and ridge extraction is integrated for avoiding loss of information.

The employed method has three parameters to be determined. σ , sth, and lth. σ is called Gaussian parameter to determine degree of smoothness in smoothing process. Sth and lth is called seeding threshold and linking threshold respectively. If absolute value of the second derivative of a pixel is more than sth, the pixel is considered as the seeding point and the line following process would begin. The extended direction of the ridge on the

pixel is calculated. Then the second derivative of a pixel that lies on the extended direction of the ridge is examined. If the absolute second derivative on the pixel is more than lth, the pixel is considered as a part of the ridge. This process will continue until there is no candidate pixel for seed point.

2.3 Line Segment Classification Stage

2.3.1 Straight Line Extraction

Road segments extracted in the previous stage are not straight line but curvature or combination of lines (poly lines) because the employed method only consider local angle relationship of each ridge pixels and does not consider global shape of segments. As the method follow pixels as long as possible, resulting segments go through two different roads or even bridge true roads and false linear object. As these mingled segments make elimination of false segments and construction of road topology difficult, curved segments must be decomposed to array of straight lines.

Another method is so called Douglas-Peucker (DP) approximation algorithm. This algorithm divides a segment at a vertex where distance from the segment becomes farthest. This simple process is recursively repeated to resulting segments until distance of all vertexes is less than a tolerance.

DP method is rather simple graph based approach. Initial approximation of a curve is the straight line that connects two end points of the curve. Then a furthest vertex from the straight line is picked up. If distance of the vertex from the straight line is more than a tolerance, this approximated poly-line is subdivided at the vertex into two shorter poly-lines. This process is repeated until every vertex on the curve is within a certain distance from resulting poly-lines.

2.3.2 False Line Elimination

Result of line detection is a blend of true road segments and false road segments such as buildings or elongated vegetated area. As these false segments often prevent correct road network construction, Classification of line segments is a necessary step. Operation of ALOS is not clear but PRISM and AVNIR-2 may not work simultaneously due to bandwidth limitation. If both images cannot be obtained simultaneously, DEM must be created from images to project both image onto proper geographic position. However, there are several difficulties in DEM creation. It is time-consuming work and DEM accuracy is insufficient in urban area that causes improper registration of imagery. For that reason, grey value image classification is worth to consider.

Multilevel threshold is expected to filter out high brightness objects such as buildings & vehicles and low brightness objects such as trees and shadows. Removal of these objects will reduce false extraction of road segments and improve correctness. This process employs automatic multilevel threshold method (ATC) proposed by Yen et al (1995). Most feature of the method is that the number of class can be determined automatically because the cost function include the number of class whereas user must specify optimal number from experience.

In the method, Cost function of this method $C(k)$ is defined as

$$C(k) = \rho(Dis(k))^{\frac{1}{2}} + (\log_2(k))^2 \quad (1)$$

where r is a positive weighting constant, k is the number of class, $Dis(k)$ is the discrepancy between after-threshold and original images. Then k and threshold value is determined to minimize $C(k)$.

For multi-spectral image, both a supervised and a unsupervised method are considered. A supervised method gives a better result than unsupervised ones if proper training data set is provided. An Unsupervised method does not need training data set but an operator must give the appropriate class number and meaningful category to a classified result. A supervised method is adequate when objects included in the image are known. An unsupervised method, to the contrary, is adequate when included objects are not known or property of each object is dispersed.

In this study, classified image is extracted using an unsupervised classification method offered by a remote sensing software, ERDAS.

2.4 Line Segment Grouping Stage

As most road segments extracted by the centre line detector are not directly connected and include some false road segments, grouping and linking of them is also needed. I have only found few articles discussing this issue. I have uses two methods proposed by Crevier (1999).

In the method, each pair is first tested by rule based screening. There are three tests for geometric relation and one for photometric relation. The angular difference between each line and the line that bridges the gap is the first inspection. The second one is the transverse gap. It measures lateral offset between lines. The third one is the longitudinal gap. It represents "net gap" without lateral offset. If any of these values exceed predetermined thresholds, this pair is rejected. The last one is the contrast difference. It checks contrast between segments and backgrounds. If a line is whiter than background whereas another line is darker than background, this pair is rejected.

Each pair that passes rule based screening test is connected and

σ	Seed threshold	Link threshold	Num. of Line	Correctness (%)	Completeness (%)
0.60	10.0	1.0	1922	10.9	69.1
0.80	10.0	1.0	1450	10.6	54.7
1.00	10.0	1.0	1174	11.9	58.2
1.20	10.0	1.0	998	11.8	47.3
1.40	8.0	0.8	952	12.8	56.1
1.60	8.0	0.8	1058	12.6	58.4
1.80	5.0	0.8	984	12.9	58.6
2.00	5.0	0.5	825	13.9	52.9
2.20	5.0	0.5	848	14.2	54.9
2.40	5.0	0.5	636	14.2	48.4
2.60	5.0	0.5	711	15.1	52.5
2.80	5.0	0.5	553	14.3	43.2
3.00	5.0	0.5	620	15.2	46.5

Table 1. Result of centre line detector.

chains of line segments are obtained. Then the validity of each chain should be examined. The measure of junction's deviation (d) of each line pair is calculated:

3. RESULT & DISCUSSIONS

3.1 Line Feature Extraction Result

According to Steger, the relationship between σ and the road width is described as follows.

$$\sigma \geq w/\sqrt{3} \tag{2}$$

Table 1 shows how many line are detected when σ is fixed. Seeding threshold and linking threshold are determined where results seem best for the human eye. Correctness constantly increases as Gaussian Kernel Parameter s become bigger. It suggests that if the kernel is too small, it picks up much noise and results in many false road segments.

Completeness, on the other hand, does not decrease constantly. It has a peak around $s=1.8$. Considering relationship between Gaussian Parameter and line width, this result suggests that width of dominating line segments in the image is less than 6 pixels (=15m in ground resolution). As many of roads in the scene are dual lane road and each lane is about 4-6m width, the result indicates $s=1.8$ may be best starting points for road centre extraction.

Table 2 shows another result of the centre line detector. Gaussian Kernel Parameter is fixed at 1,80. Change of link threshold has little effect while increase of seed threshold slightly improves correctness. This result implies that seeding threshold of 4.0 – 5.0 and linking threshold of 1.0-1.2 is best for reliable road segment extraction.

3.2 False Line Elimination Result

Table 3 shows the grey scale threshold result by ATC.. Gaussian Kernel Parameter, seeding threshold and linking threshold are set to 1.80, 5.0 and 0.5 respectively in all cases. ATC separates grey level histogram into 4 classes. Range of each class is, 22-30, 31-55, 56-88, 89-197. Then grey level of 30 and 89 are used for dark and bright thresholds.

The column “GIS road” means the total length of true road contained in the scene. This data was created manually on the GIS software, ArcView. The column “Extracted road segment” means the total road length obtained by the centre line detector. The column “Correctly extracted GIS road” indicates how many “GIS road” is extracted from the line segments obtained by the method. It is calculated as follows. At first 2.0 pixels buffer is created around the extracted road segments. Then GIS roads

Threshold (dark)	0	30	0	30
Threshold (bright)	255	255	89	89
Num. Of lines	1173	1023	1137	984
a. GIS road (pixel)	3724.9	3724.9	3724.9	3724.9
b. Extracted road segment (pixel)	22649.6	20105.6	21399.0	18859.3
c. Correctly extracted GIS road (pixel)	2284.1	2170.9	2297.1	2181.3
d. Correctly extracted road segment (pixel)	2614.9	2469.0	2580.9	2435.0
Correctness (b/d)	11.5%	12.3%	12.1%	12.9%
Completeness (a/c)	61.3%	58.3%	61.7%	58.6%

Table 3. Grey scale threshold result. Masking brighter area effectively reduces false line segments while masking darker area slightly lessens completeness.

that are within the buffer are picked up. At last, the total length of picked up GIS roads is calculated. The column “Correctly extracted road segments” indicates how many of “extracted road segments” is parallel to “GIS road”.

Then completeness can be defined as “correctly extracted GIS road” / “GIS road” and indicate how complete is the extracted road. Correctness can be defined as “correctly extracted road segment” / “extracted road segment” and indicates how the

Seed threshold	Link threshold	Num. of Line	Correctness (%)	Completeness (%)
3.0	0.8	1464	10.8	56.9
4.0	0.8	1244	12.0	61.2
5.0	0.8	1120	12.3	59.4
6.0	0.8	1018	12.1	56.5
7.0	0.8	917	13.1	56.4
5.0	0.4	1119	11.9	59.1
5.0	0.6	1119	11.9	59.1
5.0	0.8	1120	11.9	59.4
5.0	1.0	1121	11.9	60.1
5.0	1.2	1119	11.9	60.3
5.0	1.4	1118	12.0	59.9
5.0	1.6	1113	11.9	59.1

Table 2. Result of centre line detector. σ is set to 1.8 in all cases.

result is reliable.

Correctness is improved if dark and / or bright threshold are set. It suggests that introduction of multilevel threshold is reasonable to eliminate false road segments. However, completeness slightly decline if dark threshold was applied. This is because the dark threshold cuts out shadow region on roads. Shadows of buildings have elongated shapes like roads and the centre line detector recognizes these shadows as line segments. Fig.2 shows the effect of darker region filtering. Though some false line segments that go through building shadow are efficiently removed by the filtering, shadows cast on a road are also removed.

As both correctness and completeness are dissatisfying low in grey scale threshold, utilising multi-spectral images sensor is tested.

The unsupervised classification tool in ERDAS was used for the classification. Figure 2 shows a classification result. Six categories are obtained. Red, green, deep blue, light blue and grey (dark & light) represent buildings, vegetation, road, road & building and shadows respectively.

3.3 Line Grouping Result

In the rule based screening section, two characters of line pairs were checked to judge connectivity of the pairs.

For testing geometric information, three properties were inquired based on thresholds suggested in the original paper. Maximum angular difference between lines concerned and the line connecting the gap is set to 5 degree. Maximum Transverse gap, which represents how much the pair of lines is off to the side, is set to 3 pixels. Maximum Longitudinal gap, which represents how far both end points are without offset (transverse gap), is set to 1.5 times the length of shorter segment.

For testing photometric information, the author introduced “Contrast reversal test” in the original

paper. It checks contrast between lines and background. Instead of the approach, “Brightness consistency test” is adopted. Brightness of end points and brightness statistic of both segments are used. The idea behind the examination is that brightness of both segments must be similar if segments are parts of the same line. Suppose that brightness of the end point on a segment is x , mean and standard deviation of another segment is m_2 and sd_2 respectively. If the difference of x and m_2 is greater than 1.96 times of sd_2 , this line pair is rejected. The test corresponds to 10% two-tailed significance test. In the probabilistic screening section, the measure of junction’s deviation (d) is calculated for each line pairs that pass rule based screening. The pairs are sorted in ascending order of the measure and the pairs d of which is less than threshold are

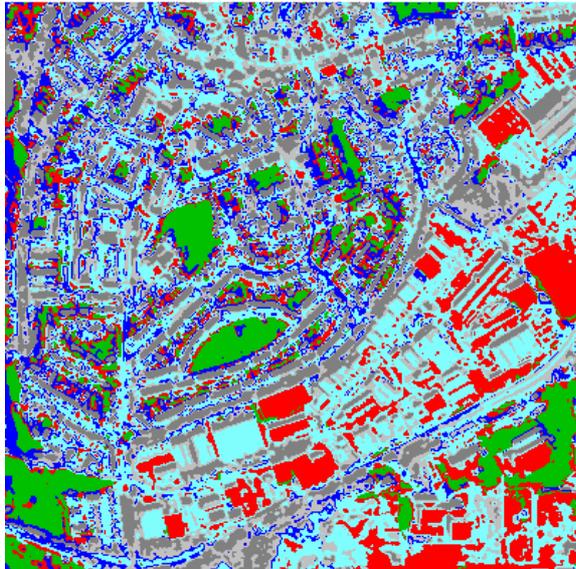


Figure 2. A classification result. All bands are used to classify image. Number of class is 15. Resulting classes are manually integrated and categorised. Red, green, deep blue, light blue and grey (dark & light) represent buildings, vegetation, road, road & building and shadows respectively.

	Case 1	Case 2	Case 3	Case 4
Min. length	5	10	5	10
Radiometric Information	No use	No use	use	use
Iteration	4	4	3	3
Num. of Lines	826	187	885	180
Correctness	13.8 %	37.5 %	13.7 %	50.5 %
Completeness	55.4 %	36.2 %	49.7%	32.5 %

Table 4. Result of grouping. Simply minified image was used and same set of parameter was applied for all cases in the edge extracting stage. Limitation of minimum segments length in grouping stage sufficiently improve correctness though it reduces completeness. In case 4, improvement of correctness is significant.

connected. In the original paper, probability of chains are experimentally calculated and used for validating segment connectivity to avoid hard threshold value. In our study, however, hard threshold method is applied to check effectiveness of grouping process.

Table 4 shows a result of grouping process. The simply minified image was used. Parameters for the edge extraction

process are 1.0 for sigma, 7.0 for edge seeding threshold and 0.5 for edge linking threshold. 1108 segments are extracted as a result.

In all cases, segments longer than 5 pixels are used for next iteration. In case 1, geometric relation only is considered to investigate line connectivity. The process converges at 4 iterations. In case 2, the segments more than 10 pixels are picked up and evaluated. In case 3 & 4, photometric relation is also considered. It is very clear that cut-off of small segments considerably improves correctness. If photometric information is also considered in grouping process, degree of improvement in correctness becomes greater. In case 4, where photometric information is used in grouping process and small segments are deleted after convergence, correctness becomes 50% though completeness drops to 30%.

Table 6 shows another result of grouping process. The original image (Figure 1) is used. Parameters for the edge extraction are 1.8, 4.0 1.0 for sigma, seeding threshold and linking threshold respectively. The segments shorter than 7.0 pixels are not accepted. 1433 lines are extracted as the grouping process.

4. CONCLUSION

A road extraction method from ALOS PRISM image have been investigated and tested. This process consists of three stages; pre processing stage, edge extraction stage and grouping stage.

In pre processing stage, an automatic brightness threshold method were applied though elimination of brighter / darker area slightly improved final result.

In edge extraction stage, a line centre diction method has been introduced and evaluated. Better combination of parameters has been searched but correctness was very poor. It suggests the need of some filtering method to distinguish true road segments from false ones. Some minified images have been also tested but their correctness and completeness were as same as that of original image.

In grouping stage, use of segment length and radiometric property condition significantly improved correctness though completeness drops by 50%.

As a whole, more improvement is needed to rise both of correctness and completeness. One idea to achieve it is excluding building area and vegetation area in pre-processing stage. Another idea is using GIS road data as a guide and classifies extracted road candidates. Hu and Tao (2002) use the concept of saliency. They applied segment grouping to “the most salient” lines and later “less salient” lines were added and grouping process was continued. This method might be worth to consider for this study.

5. REFERENCES

Baltsavias, E.P., 2002. Object Extraction and Revision by Image Analysis Using Existing Geospatial Data and Knowledge: State-of-the-art and Steps Towards Operational Systems. IAPRS vol. 34, Part 2, pp. 13-22.

Baumgartner, A., Steger, Mayer, C., Eckstein, H. W., and Ebner, H., 1999. Automatic Road Extraction Based on Multi-Scale, Grouping, and Context. PE&RS, 65(7), pp. 777 – 785.

Crevier, D., 1999. A Probabilistic Method for Extracting Chains of Collinear Segments. Computer Vision and Image Understanding, 76(1), pp. 36-53.

Hu, X., and Tao, C. V., 2002. Automatic Main Road Extraction From High Resolution Satellite Imagery. IAPRS, vol. 34, part 2, pp. 203-208.

Park, S., and Kim, T., 2001. Semi-Automatic Road Extraction Algorithm from IKONOS Images Using Template Matching. Proceeding of the 22nd ACRS.

Price, K., 1999. Road Grid Extraction and Verification. IAPRS, vol.32, Part 3-2W5, pp. 101-106.

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

Treash, K., and Amaratunga, K., 2000. Automatic Road Detection in Grayscale Aerial Images. ASCE Journal of Computing in Civil Engineering, 14(1), pp. 60-69.

Zhao, H., Kumagai, J., Nakagawa, M., and Shibasaki, R., 2002. Semi-automatic Road Extraction From High-resolution Satellite Image. Proceedings of Photogrammetric Computer Vision, A, pp. 406-411.

ACKNOWLEDGEMENT

The author would like to thank Ministry of Education, Culture, Sports, Science and Technology for provide a scholarship. The author also greatly appreciates Prof. I. Dowman and staff & member of the University College London for making opportunity to study as a honorary research fellow and to have useful discussions.