
© Copyright 2009 [please consult the authors]
ENHANCING DIGITAL ROAD MAP WITH LANE DETAILS EXTRACTED FROM LARGE-SCALE STEREO AERIAL IMAGERY USING OBJECT-ORIENTED IMAGE ANALYSIS

Hang Jin, Zhengrong Li, and Yanming Feng
School of Information Technology, Queensland University of Technology, Brisbane, Queensland 4101, Australia
jinhang3695@gmail.com

ABSTRACT

Precise, up-to-date and increasingly detailed road maps are crucial for various advanced road applications, such as lane-level vehicle navigation, and advanced driver assistant systems. With the very high resolution (VHR) imagery from digital airborne sources, it will greatly facilitate the data acquisition, data collection and updates if the road details can be automatically extracted from the aerial images. In this paper, we proposed an effective approach to detect road lane information from aerial images with employment of the object-oriented image analysis method. Our proposed algorithm starts with constructing the DSM and true orthophotos from the stereo images. The road lane details are detected using an object-oriented rule based image classification approach. Due to the affection of other objects with similar spectral and geometrical attributes, the extracted road lanes are filtered with the road surface obtained by a progressive two-class decision classifier. The generated road network is evaluated using the datasets provided by Queensland department of Main Roads. The evaluation shows completeness values that range between 76% and 98% and correctness values that range between 82% and 97%.

INTRODUCTION

Accurate and detailed lane-based road features could benefit a large number of transportation and advanced vehicle navigation applications, which includes lane departure warning and lane-level navigation. Lane-based vehicle navigation could be employed to assist drivers in complicated roadway networks, particularly during congested conditions. However, lane information is not readily available and often must be acquired for the areas of interest. A common practice of acquiring lane-based roadway data, using the probe vehicle equipped with carrier-phase DGPS travelling down different lanes (Du and Barth, 2008), is somewhat tedious. Other drawbacks include driver variations within lanes and lane changes, as well as the fact that lane boundary demarcations cannot be accurately determined. Another method for obtaining lane-level road network data is through the use of high resolution images. As manual extraction of road details from imagery is very time consuming, automated methods have the potential to improve the speed and utility for road mapping and are therefore highly desirable.

The extraction of roads information from digital images has drawn considerable attention in the last few decades (Amini et al., 2002). Early approaches for road details detection were developed using both low and high-resolution aerial photographs (Heipke et al., 1995, Mayer et al., 1997, Gong and Wang, 1997). In high-resolution aerial images, geometrical entities, such as structure and shape, play a crucial role in road recognition. Accordingly roads are often modelled as continuous and elongated homogenous regions with nearly constant width. Many road extraction algorithms are based on the analysis of geometrical and photometrical attributes. For instance, Amini (2002) presented an object-based approach for automatic extraction of main roads in large scale images, in which the road sides are extracted by combining straight lines and road skeleton extraction. Renaud (2004) reconstructed road model by extracting the road segments with parallel sides and intersections respectively, where an algorithm called “DoubleSnakes” has been proposed to extract road parallel lines. In (Baumgartner et al., 1997) the anti-parallel linear edge segments are extracted and the road segments are selected, which are then linked to extract roads with a multi-resolution approach.
An existing road database provides not only the approximations but also other usual information, such as road attributes, global context and so on, referring to the descriptions in (Zhang, 2004, Baltsavias and Zhang, 2005). In addition, other available geodata such as Digital Surface Model (DSM) and Digital Terrain Model (DTM) data can also be incorporated to provide complimentary and sufficient redundant information for benefiting the extraction processing and thus improving the results. In this regard, Hinz (2003, 2004) integrated the knowledge of roads and their context using explicitly formulated scale-dependent models to extract the detailed road information in urban area. In addition, an integrated system for automatic mapping of urban and suburban roads from high-resolution satellite imagery is found in (Jin and Davis, 2005). Multi-source data fusion effectively enriches and improves road information. However, these existing road detection approaches are basically limited to the extraction of the road level data rather than the lane level data. Although the detection of road lanes and road sides was also suggested (Zhang, 2004, Baltsavias and Zhang, 2005, Hinz, 2004, Hinz and Baumgartner, 2003) to provide additional cues for road extraction, they were based on line detection and would be greatly affected by noises within large scale aerial imagery. The rich information content provided by VHR imagery of high quality can dramatically aggravates the process of pixel labeling, and the available state-of-the-art image analysis procedures – basically pixel-based approaches – have considerable difficulties in dealing with the rich information content (Blaschke and Strobl, 2001). In this paper, we will design an approach based on object-oriented image analysis to automatically detect the road lane information and further reconstruct the lane level digital road map. The rest of the paper is organized as follows. Firstly, preprocessing and orthorectification procedure of the original aerial image are described so as to remove the displacement caused by the tilt of the sensor and terrain relief. Then, we detect the road lane details using an object-oriented rule based image classification. After that, road surface, which is extracted by using a progressive decision classifier, is further utilized to filter out misclassified objects that have the similar spectral and geometrical feature as lane data. The extracted lane data are further vectorized and smoothed. And finally, the quality evaluation for the extracted road lane is presented. Besides, the concluding remarks and the research findings are summarized.

**PREPROCESSING**

The purpose of preprocessing is to improve the quality of the image and gather the data for the further processing. To use the aerial images for obtaining road lane details, the images have to be georectified, since only the features extracted from orthoimages can correctly represent the position of the ground objects. An orthoimage is the aerial image which is geometrically corrected (“orthorectified”) such that its scale is uniform, namely, the image has the same scale of distortion as a map (Mori, 1997). As certain attributes of objects to be extracted, such as the length and width of a road, have known, road bounded values can be directly compared to the attribute values of the features that are extracted in the orthoimage (Baltsavias, 1996).

![Flow Chart](chart.png)

**Fig.1:** The flow chart of ortho-image production (Mori, 1997)

The procedure for geometric processing of stereo images includes two main steps: (i) to measure the ground control points and the image tie points; and (ii) to triangulate the images using polynomial approximation for the exterior orientation parameters. The basic workflow is shown in Fig.1. The DSM can be computed using a pyramid-based crossing-correlation and subsequent relaxation matching techniques such as implemented in...
the Virtuozo digital photogrammetric station. The DSM includes modelling of the surface of all entities such as buildings, trees, and not just the terrain. Thus, the generated DSM was edited manually to remove the ground objects so as to obtain the Digital Terrain Model (DTM), see, e.g., Fig.2 (a), which will be used to geometrically correct the raw image. The orthorectified image is shown in Fig.2 (b).

![Fig.2: (a) DTM, and (b) The orthorectified image](image)

The raw aerial images typically have a low local contrast due to a wide radiometric dynamic range of the scene content and possible atmospheric disturbances. A typical global contrast enhancement (e.g., Linear, Normalized) cannot simultaneously produce good local contrast at both ends of the brightness range. Therefore, Wallis adaptive filter (Baltsavias et al., 2001), which adjusts brightness values in local areas so that the local mean and standard deviation match user-specified target values, is further used to enhance the edge features, thus facilitates the image segmentation and lane extraction process. The image is firstly transformed using Principal Component Analysis (PCA), and the Wallis filter is then applied on the first component to stretch the local contrast, after that the inverse principal component is calculated.

![Fig.3: Work flow of the proposed approach](image)

After the preprocessing, the orthorectified image can be further utilized to extract road lane information. As can be seen from Fig.3, the proposed approach mainly includes: i) lane feature extraction, and ii) the lane feature filter, which will be further depicted in the next two sections.
LANE EXTRACTION

With the availability of VHR imagery, object-based analysis has recently attracted extensive interest in a variety of remote sensing applications, as it provides specific advantages and unique capabilities over traditional pixel-based analysis (Jin and Paswaters, 2007), such as it can overcome the H-resolution problem and salt-and-pepper effect. Therefore, an object-oriented image classification method is utilized for the lane information extraction in our approach.

A prerequisite to classification is image segmentation, which is the subdivision of an image into separated regions. Image objects resulting from segmentation represent image object primitives, serving as information carriers and building blocks for further classification (Baatz et al., 2004). In the rest of the section, we give two components of the object-oriented model: image segmentation and object-based classification.

Image segmentation

Image segmentation is usually the first step in the bottom-up image analysis strategy, where the success of the image segmentation is critical to higher-level tasks such as feature extraction and object recognition. The purpose is to group individual pixels into homogeneous regions to work as the basic analysis element. By this way, spatial, contextual and semantic information was able to be extracted and facilitates the object-based analysis.

The image segmentation is actually a bottom-up region merging process – grouping adjacent pixels by the means of region growing technique. It starts with one pixel shaping one image object or region, and the merge continues until a user specified criteria is reached. The merge decision is based on local homogeneity criteria. The algorithm guarantees a regular spatial distribution of treated image objects. The stop criterion for the region-merging process is given by the parameter “scale”, which determines the maximum allowed overall heterogeneity of the segments (Benz et al., 2004). Fig. 4 (a) shows the Wallis filtered image, while the image segmentation result is as illustrated in Fig.4 (b). The lane lines are presented with white elongated features in the segmented image.

![Image](image.png)

Fig.4: (a) Wallis filtered image, and (b) Image segmentation result

Object-based classification

After segmentation, the spectral and spatial attributes can be calculated on a per region base. The attributes include (Russ, 2006):

- **Spectral attributes**: mean, minimum, maximum of intensity values on each spectral band, and band ratio on each spectral band.
- **Shape attributes**: shape descriptors of regions such as area, perimeter, compactness, elongation, major length and minor length.

In remote sensing data, the uncertainties are inherent in the whole information extraction system. Thus fuzzy classification, which approximates human knowledge and reasoning process, is employed to handle the uncertainties which are inherent in image data and human knowledge. Each class of a classification scheme contains a class description. Each class description consists of a set of fuzzy expressions allowing the evaluation of specific features and their logical operation.
### Tab.1: Attributes used in our lane extraction process

<table>
<thead>
<tr>
<th>Membership function</th>
<th>Description</th>
<th>Road lane</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average band of x</td>
<td>Average value of the pixels comprising the region in band x (R, G and B)</td>
<td>Relatively large</td>
</tr>
<tr>
<td>NDVI</td>
<td>A normalized band ratio is computed between R and IR band, we use R and G band instead</td>
<td>Low</td>
</tr>
<tr>
<td>Major axis</td>
<td>The length of the major axis of an oriented bounding box enclosing the polygon</td>
<td>Large</td>
</tr>
<tr>
<td>Minor axis</td>
<td>The length of the minor axis of an oriented bounding box enclosing the polygon</td>
<td>Small</td>
</tr>
<tr>
<td>Main direction</td>
<td>The angle subtended by the major axis of the polygon and the x-axis in degrees</td>
<td>Almost constant</td>
</tr>
<tr>
<td>Elongation</td>
<td>A shape measure that indicates the ratio of the major axis of the polygon to the minor axis of the polygon</td>
<td>Large</td>
</tr>
<tr>
<td>Area</td>
<td>Total areas of the polygon, minus the area of the holes</td>
<td>In a certain range</td>
</tr>
</tbody>
</table>

The classification of image objects can be performed by using membership functions, based on fuzzy logic theory combined with user-defined rules. Spectral, shape, and statistical characteristics as well as relationships between linked levels of the image objects can be used in the rule base to combine objects into meaningful classes (Benz et al., 2004).

![Fig.5: Detected road lane features](image)

Road lane lines are classified by the membership functions as shown in Tab.1. We take NDVI for example, the NDVI of road lanes should be relatively low, which means objects with NDVI values lower than a thresholding are classified as road marks. All the functions listed in Tab.1 generate equal weight in the analysis and were combined by a logic “and” (minimum operator) function, which equals the minimum fulfillment of the single statements. The result generated by the rule-based classification is presented in Fig.5. As we can see that, almost all road lanes are correctly extracted from the imagery. However, two bare land objects (shown in the circles), which have the similar spatial and spectral attributes as road lane features, are also misclassified into the lane class. Therefore, we will employ the road surface to further filter out these misclassified features.

**LANE FILTER**

In order to appropriately extract the road surface object, we must firstly understand how a road’s physical characteristics influence its visual characteristics (Fortier et al., 2001). On remote sensing imagery, roads are recognizable from their distinctive tone, texture and shape. Consisting of the same paving material, roads have a uniform tone as well as a smooth and identical texture. There is a drastic change in tone and texture across road boundaries. We concluded the road features as following:

- The steepness of road has an upper bound, thus slope can be utilized to remove objects that have a great gradient change, such as trees and houses.
- Roads have little or zero normalized difference vegetation index (NDVI) value, so NDVI can be employed to filter out vegetation from the scene.
- The road surfaces are relatively dark, so they should have a low average pixel value, and it is also likely to vary only slowly.
- Road surface has a relatively large area, thus objects that have a smaller area can be further sieved.

![Progressive decision classifier for road surface](image)

**Fig.6:** Progressive decision classifier for road surface

A decision tree classifier as shown in Fig.6 is utilized for the extraction of road surface. The decision tree classifier (Jia and Richards, 1998) performs multistage classification by using a series of binary decisions to place pixels into classes. Each decision divides the pixels in a set of images into two classes based on an expression. The result of the decision classification is illustrated in Fig.7 (a). As we can find that many spots objects are misclassified into road surface class, and the classified road surfaces are also quit disintegrated. Therefore, an area filter is firstly used to remove these noise points, and then morphological closing is employed to smooth the road surface. The final road surface is presented in Fig.7 (b).

![Classification result](image) ![Extracted road surface](image)

**Fig.7:** (a) Classification result, and (b) Extracted road surface

Up to now, we have obtained the classified result of road surface, based on which the lane data obtained in section 2 can be further filtered accordingly. The objects located outside the road surface are sieved. After the lane detail filtering operation, we also employ the thinning algorithm proposed by Wang and Zhang (Wang and Zhang, 1989) to thin the extracted lane features to a pixel width and a road network pruning method proposed by Jin (Jin et al., 2008) to vectorize the lane lines and remove the short dangling branches of the centreline caused by the thinning process. The Douglas-Peucker simplification (Richards and Jia, 2006) and
Bezier interpolation methods are finally utilized to simplify and smooth the lane lines. The vectorized lane lines are illustrated in Fig. 8 (a).

![Fig. 8: (a) Vectorized lane features, and (b) Disconnected lane segments](image)

Fig. 8 (b) is the zoomed rectangular area in Fig 8 (a). It can be easily found that the border lane line is broken into two segments. Therefore, a line linking algorithm presented in (Amini et al., 2002) is employed to link the lane segments by using the geometric properties. The linking algorithm consists of a search for a pair of lines that satisfy the colinearity and proximity relations that make them candidates for grouping. Four constraints are used: linking radius, colinearity, endpoints, and overlap, they are presented as: the linking segment candidates must intersect the circular area with a defined radius; they also must be approximately collinear; and the endpoints of them must be close; finally they must not overlap too much. The detailed information can be referred to (Amini et al., 2002). The final linked lanes are presented in Fig. 9 (a), and Fig. 9 (b) shows the generated road lane information overlapped on the raw aerial imagery, which fit the original lanes very well.

![Fig. 9: (a) Linked road lane segments, and (b) Final lane data overlapped on the raw imagery](image)

RESULT AND EVALUATION

The proposed road lane extraction strategy is tested using the aerial images of Gympie, Queensland, acquired on 28 Nov, 2008. The detail information about the image dataset is presented in Tab. 2.

**Tab. 2: Image specifications of Gympie dataset**

<table>
<thead>
<tr>
<th>Specification</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of color channels</td>
<td>3, RGB</td>
</tr>
<tr>
<td>Scan resolution</td>
<td>9 μm</td>
</tr>
<tr>
<td>Camera focal length</td>
<td>105.2 mm</td>
</tr>
<tr>
<td>Flying height</td>
<td>~1000 m</td>
</tr>
<tr>
<td>Image scale</td>
<td>1:11081</td>
</tr>
<tr>
<td>Ground resolution</td>
<td>~0.1 m</td>
</tr>
</tbody>
</table>
Forward overlap | ~75%/
---|---
Orientation | Known

Within the test area, the road lane lines were firstly manually delineated. The manually delineated lane data was used as a reference road to assess the accuracy of the automatic lane extraction algorithm outlined in Sections 2, 3 and 4. Three evaluation indicators presented in reference (Wang and Liu, 1994) are calculated for the quantitative comparison.

Let $N_{ce}$, $N_{tr}$ and $N_{te}$ be the number of correctly extracted road lane pixels, the true lane pixels and the total number of extracted lane pixels, respectively, the evaluation indicators are computed as

\[ \text{overall accuracy} = \frac{N_{ce}}{N_{tr}} \]  
\[ \text{commission error} = \frac{(N_{te} - N_{ce})}{N_{tr}} \]  
\[ \text{omission error} = 1 - \frac{N_{ce}}{N_{tr}} \]

Four representative test areas with a size about 2 km$^2$ were selected from the dataset. For all these test sites, the overall accuracy ranges between 76% and 98%, the commission error between 3% and 18%, and the omission error between 2% and 24%. The correctly extracted lane lines are greatly affected by the trees over the road and the shadows casted by the trees along the road. However, the results can be further improved by using knowledge based image analysis combining parallel lines analysis, DSM and vegetation recognition.

**CONCLUSION**

This study has demonstrated the potential use of the objects-oriented approach as a tool for effectively mapping road lane information with high resolution aerial imagery. The spatial comparison of the road lane data generated by the proposed approach with manually delineated lanes have shown a very high degree of correctness and completeness despite that some are missing or distorted slightly. Overall, certain success has been achieved in our attempt of extraction of road lane information from high-resolution stereo aerial imagery based on image analysis. Further improvement may be achieved by using knowledge based image analysis combining parallel lines analysis, DSM and vegetation recognition.

**Acknowledgements**

This work is partially supported by Chinese Scholarship Council (Grant No: 2007101585). The authors are grateful to the Department of Transport and Main roads, Queensland for providing the testing aerial image dataset. Our sincere thanks go to Mr Bofeng Li for his good suggestions in the data processing.

**REFERENCES**


