# ROBUST DETECTION OF ROAD JUNCTIONS IN VHR IMAGES USING AN IMPROVED RIDGE DETECTOR

S.Gautama, W.Goeman, J.D'Haeyer

Dept. Telecommunication and Information Processing, Ghent University
St.Pietersnieuwstraat 41, B-9000 Gent, Belgium
sidharta.gautama@ugent.be

**KEY WORDS:** GIS, object recognition, road detection, change detection

**ABSTRACT:**

In this paper we present a novel technique for automatic detection of road junctions in VHR satellite images. The detection is based on an initially detected road network using a differential ridge detector. This is a standard technique for road detection. However, the performance of the ridge detector degrades in the vicinity of junctions because the line model on which it is based does not hold anymore. We analyze the content and quality of the derived information layers of the ridge detector and show which information is useful for the detection of junctions. The detected network is improved using a region growing and thinning strategy. Junctions are detected based on this network using a shape analysis. This analysis puts restrictions on the appearance of junctions and allows for an efficient filtering of false alarms. Experimental results performed on IKONOS images over the city of Ghent show a reasonable detection rate with a very low false alarm rate can be achieved. This low false alarm rate is important for our purpose of quality assessment as it requires reliable image information to make a robust comparison with the road database.

## 1. INTRODUCTION

A major challenge in the production and use of geographic information is assessment and control of the quality of the spatial data. The rapid growing number of sources of geospatial data, ranging from high-resolution satellite and airborne sensors, GPS, and derivative geospatial products, pose severe problems for integrating data. Content providers face the problem of continuously ensuring that the information they produce is reliable, accurate and up-to-date. Integrity constraints are able to resolve certain issues in the data, like valid attribute values or relationships between data objects. The main issue is however the consistency of the data with respect to the current "real-world" situation. Today the industry still relies on human operators, who collect and interpret aerial photos and field data to check and correct the current state of the data. Especially for detailed data over large regions, like digital road maps or topographic maps, this is a very labour-intensive and costly process. In addition, human processing is a source of error and inconsistency. Automated detection of change and anomalies in the existing databases using image information can form an essential tool to support quality control and maintenance of spatial information.

The main problem however are the differences in data representation. To be able to compare geospatial vector data with images, the information in the images needs to be described in terms of object features. Automatic detection of man-made objects is a difficult problem. Shadow, occlusion and variety in appearance all give rise to a fragmented and imprecise description of the image content, especially if consistent detection is required over large datasets. Within the field of automatic quality assessment, there is a high need for powerful detection techniques but with a strong emphasis on reliability. From the viewpoint of the data provider, a statement about the quality of his data is only useful if the statement can be made with high reliability.

In our work, we investigate a system for change detection based on object based spatial registration, where detected object features in the image are registered to corresponding features in the vector data. The system consists out of two stages: 1) a low-level feature detection process, which extracts information about the road network, and 2) a high-level matching process, which uses graph matching to find correspondences between the detected image information and the road vector data (cfr.Figure 1). The graph matching process is driven by the spatial relations between the features and takes into account different errors that can occur (e.g. spatial inaccuracy, data inconsistencies between image and vector data). The matched features can be used to calculate a local transformation between image and vector data, which is able to compensate for local distortions that can occur between the datasets. Additionally the object-to-object mapping is useful to define measures of change between datasets.

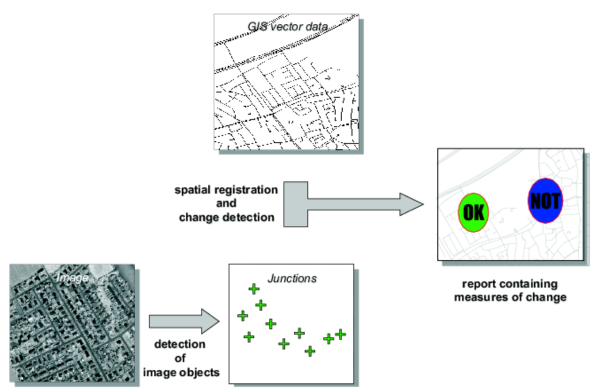



Figure 1. System overview for object based quality assessment.

Although much effort has been spent on designing algorithms for road detection, the complexity of the problem is still not fully tackled. In this paper we examine road detection using a ridge detector. Lines in an image can be seen as narrow valleys or ridges in the intensity surface if one views the image as a terrain model. This ridge model works well for roads which are bordered by homogeneous regions like fields. However in the more general case, it often leads to more fragmented results due to disturbances like cars, shadows, trees etc. which perturbate the ridge structure. In addition, there is an added difficulty when using the detected road segments for performing change detection on a road vector layer. The road segments from the image are represented by connected chains of pixels, whereas the vector data consists out of polylines. Moreover the chains can be fragmented or contain small disturbances which make a good comparison with the polylines difficult. Therefore, as a first phase in our work we concentrate on the detection of road junctions.

## 2. ROAD DETECTION

Lines in an image can be seen as narrow valleys or ridges in the intensity surface if one views the image as a terrain model. Steger (1998) reviewed different approaches to line detection. In this work, line detection is performed based on polynomial interpolation to determine pixels belonging to road structures in the image, the "facet model" (cfr. Haralick and Watson, 1981). This is a standard method for ridge detection. The image is regarded as a function *I(i,j)*. Lines are detected as ridges and ravines in this function by localy approximating the image function by its second order Taylor polynomial. The polynomial is used to approximate first and second order derivatives of the image function in each pixel. The direction of the line can be determined from the Hessian matrix of the Taylor polynomial. The gradient and curvature information in each pixel are used to classify a pixel in a number of topological classes based on their sign or magnitude. Line points are mainly characterized by a high second directional derivative, i.e. a high curvature perpendicular to the line direction.

The calculation of the partial derivatives can be done in various ways. The facet model determines a least squares fit of a polynomial $F$ to the image data $I$ over a window of size $N=w^2$ with window size $w$. The origin is chosen in the central pixel of the window. The value of the polynomial $F$ in pixel *(i,j)* is given by:

$$\begin{aligned} F(i,j,\vec{\theta}) &= a_1 + a_2 i + a_3 j + a_4 i^2 + a_5 ij + a_6 j^2 \\ &= \vec{m}^T \vec{\theta} \\ \vec{m} &= \begin{bmatrix} 1 & i & j & i^2 & ij & j^2 \end{bmatrix}^T \\ \vec{\theta} &= \begin{bmatrix} a_1 & \cdots & a_6 \end{bmatrix}^T \end{aligned} \tag{1}$$

The facet model for line detection searches the least-squares solution $\vec{\theta}$, given the image data $\vec{x}$ containing the intensity value *I(i,j)* in each pixel *(i,j)*:

$$\begin{aligned} &\arg\min_{\vec{\theta}} r(\vec{\theta}) \text{ with } r(\vec{\theta}) = \left\| M\vec{\theta} - \vec{x} \right\|^2 \\ &M = \begin{bmatrix} 1 & i_1 & j_1 & i_1^2 & i_1 j_1 & j_1^2 \\ \vdots & & & & & \vdots \\ 1 & i_N & j_N & i_N^2 & i_N j_N & j_N^2 \end{bmatrix} \in \mathbb{Z}^{N\times 6} \\ &\vec{x} = \begin{bmatrix} I(i_1,j_1) & \ldots & I(i_N,j_N) \end{bmatrix}^T \in \mathbb{R}^{N\times 1} \end{aligned} \tag{2}$$

where $w_{1/2} = \lfloor w/2 \rfloor$.

This leads to the linear system $M^T M \vec{\theta} = M^T \vec{x}$ with the solution $\vec{\theta}_0$ given by $\vec{\theta}_0 = (M^T M)^{-1} M^T \vec{x}$. The matrix $M$ is independent of the position of the window within the image, meaning that the calculation of $(M^T M)^{-1} M^T$ needs to be

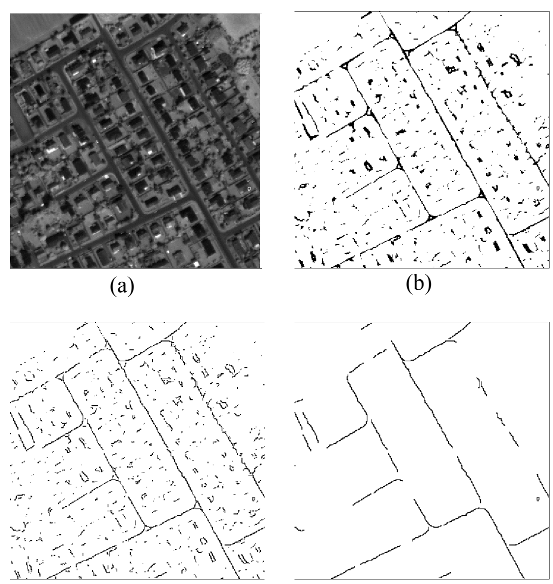
| $\|G\|$ | $\lambda_1$ | $\lambda_2$ | class |
|---|---|---|---|
| 0 | 0 | 0 | flat |
| 0 | - | - | peak |
| 0 | + | + | valley |
| 0 | - | 0 | ridge |
| 0 | + | 0 | valley |
| + | 0 | 0 | slope |
| + | - | - | slope |
| + | + | + | slope |

Table 1. Classification of the image structure based on gradient and eigenvalues of the Hessian.

performed only once for the processing of an image with a fixed window size $w$. On the basis of the parameters $\vec{\theta}$ of the interpolated surface $F$, the gradient and Hessian in a certain pixel can be calculated:

$$\begin{aligned} gradient(I) &= \left[ \frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right]^T = \begin{bmatrix} a_2 + 2a_4 i + a_5 j \\ a_3 + a_5 i + 2a_6 j \end{bmatrix} \\ Hessian(I) &= \begin{bmatrix} \frac{\partial^2 I}{\partial x^2} & \frac{\partial^2 I}{\partial x \partial y} \\ \frac{\partial^2 I}{\partial y \partial x} & \frac{\partial^2 I}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2a_4 & a_5 \\ a_5 & 2a_6 \end{bmatrix} \end{aligned} \tag{3}$$

Based on the gradient and eigenvalues of the Hessian, each pixel in the image can be assigned a topological class based on the sign and magnitude of the gradient and eigenvalues (cfr. Table 1). For road detection, we are interested in the ridge and valley class. Figure 2 shows an example of ridge detection performed on an extract of a IKONOS panchromatic scene above Ghent. Detection is performed using a window size *w=9*. A morphological thinning operator is performed on the raw detection result to produce lines of single pixel width (cfr. Arcelli et.al., 1981).



(a) (b) (c) (d)

Figure 2. Example of ridge detection. (a) original, (b) detected ridge pixels, (c) thinned result, (d) filter on minimal segment length.

## 3. JUNCTION DETECTION

The roads that can be extracted using the ridge detector are not of sufficient quality to be useful for registration with a road vector layer. The main difficulty is the difference in representation between the pixel chains that are detected in the image and the polyline vectors that represent roads in the database. This difference hinders the correspondence problem considerably. A much more robust registration object is necessary. Road junctions are good candidates since in their abstract form, they can be represented as point objects both in the image as well as in the database. This reduces change detection to the comparison of sets of points, for which several reliable techniques are available (e.g. Gautama and Borghgraef, 2003).

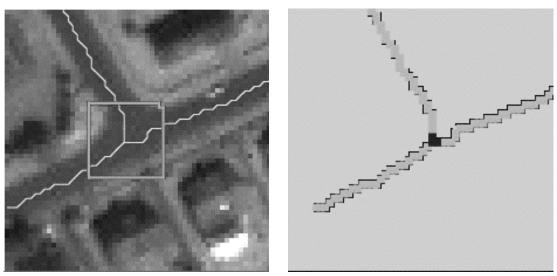


Figure 3. A junction defined a chain pixel with three or more neighbours.

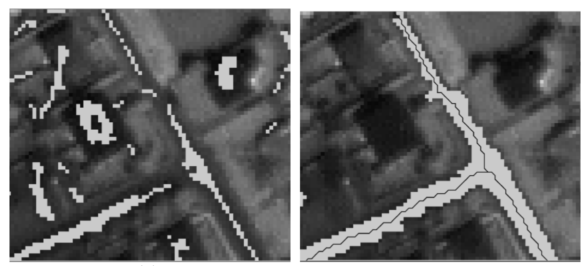
We model a road junctions as points in the road network, at which three or more road segments meet. A similar definition has been used by Wiedemann (2002). This means that built upon the road network that is detected using ridge detection, we look for pixels in the pixel chains which have three or more neighbours. We chose this strategy above corner detection, because corner detection gives many spurious responses not belonging to road junctions, which are not easy to filter out. Our method is more specifically tuned to the road network logic.

A major problem however is that the road network as is detected, fails in the vicinity of junctions since the ridge model does not hold anymore. At the junction, the intensity surface will not appear as a valley or a ridge but as a flat spot. Pixels at a junction will show a low gradient and a low curvature in both directions. Figure 6 illustrates this problem. Figure 6b shows a junction with the detected pixel chains in overlay. The road network is typically broken at junctions. Figure 6a shows the corresponding eigenvalue $\lambda_1$. In the center of the junction, the flat spot area can be seen quite clearly. It should be noted however that this flat spot not only occurs at a junction, but it can also be caused by buildings and other compact structures. A flat spot as such is therefore not sufficient to reliably detect junctions and the information about road network should be used to further characterize a junction.

For this, we implemented a region growing scheme which extends the initial road segments with regions which show a similar grey value. Region growing is a standard segmentation algorithm which works with either grayscale or multispectral images (cfr. Levine and Shaheen, 1981). It is a queue-based algorithm which iteratively grows regions starting from seed regions to form homogeneous area s in the image. In the simplest case, a pixel will be included in a region if it is adjacent to another pixel in the region that has a intensity value that differs less than a given threshold *T*. If there are multiple bands, the threshold criterium must hold for all spectral bands. The algorithm can be extended by applying adaptive thresholding. The threshold is modified dynamically according to the mean and standard deviation of the region as it is being grown. The modification equation is based on an algorithm by Levine and Shaheen (1981) and is given by:

$$T^{t+1} = \left(1 - \min(0.8, \frac{\sigma_{region}}{\mu_{region}})\right).T^t \tag{4}$$

with $T^0=T$. Thus the adaptive threshold will never be larger than the value of the initial threshold *T*, but it can become much smaller. Using the adaptive threshold can help to prevent "bleeding" across slow image gradients.



(a) (b)

Figure 4. (a) initial detection,
(b) detection after region growing and thinning.

The process to improve the initial ridge detection result contains the following steps:

1. the road segments that have been detected using ridge detected are filtered on size, where only segments of a certain size are kept as initial seeds;
2. based on these seeds, region growing is applied using the adaptive threshold;
3. morphological thinning is performed to produce a line of single pixel width.

In the last step, the maximum supression technique which is typically used in road detection to produce single pixel lines, cannot be applied because of the flat spot that occurs in junctions. Selecting the pixels of maximum curvature would in this case produce unwanted cycles around the flat spot. The thinning process does not have this problem and can produce cleaner junctions.

For roads which are adequately detected, this process proves to be sufficient in many cases to bridge the bad spots at junctions. Figure 4 shows the result after region growing. Figure 4a shows the initially detected ridge pixels. Figure 4b shows the resulting segments after region growing and the pixel chains after thinning. Segments of small size have been filtered out.

Based on the improved road network, road junctions can be detected using the neighbour definition. The simple scheme is of course not fool proof. A cheap and efficient verification to filter out false alarms is to check if in the vicinity of a hypothetical road junction a flat spot exists. Road junctions are seen by the ridge detector as a flat region. By simply applying different thresholds on the gradient and the curvature, which have already been calculated for the detection of lines, we can detect the flat spots and verify our crossroads. Figure 5 shows an example of crossroad detection. The boxes are detected road junctions using the neighbour model. Crosses are road junctions which are verified and retained using the flat spot criterium.

Although in this paper the experiments are restricted to the panchromatic image, the technique can be applied on the multispectral image as well.

Figure 5. Detected road junctions. Boxes show initial detection. Crosses show detection after filtering with flat spot criterium.

## 4. EXPERIMENTAL RESULTS

We applied our technique on an IKONOS panchromatic scene above Ghent. The high resolution of this sensor makes it not evident that standard models for junction detection hold.

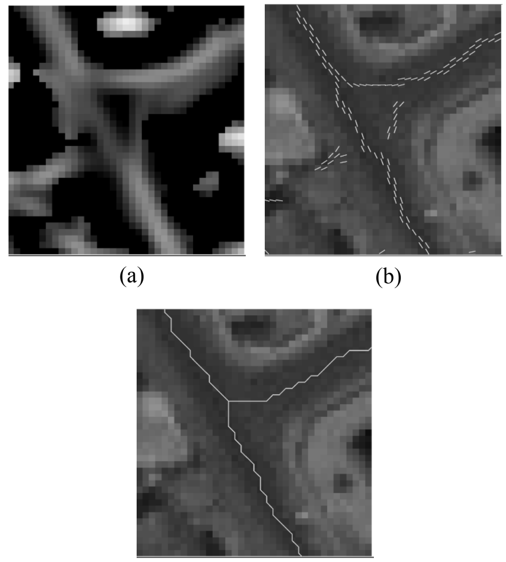
### 4.1 Qualitative discussion of different methodologies

We start our discussion by comparing the information content that is used by two techniques on junction detection that are proposed in the literature. Specific in our discussion is the emphasis on very-high-resolution imagery (<1m). Most of the techniques have been derived for resolutions where roads appear as lines of small width. Therefore, the properties of the junctions differ going from high to very-high resolution and it is not straightforward to extrapolate the published results.

Beaudet (1978) proposed a rotationally invariant operator named DET which corresponds to the determinant of the Hessian matrix. He suggested that junction detection can be done by thresholding the absolute values at the extrema of this operator. If we look at the example in Fig.6a where the respons of the maximum eigenvalue is plotted, we see that the junction appears a "flat spot" where the eigenvalue has a low value. The main problem with this condition is that in very-high-resolution, many other structures show the same property (e.g. houses). This makes it difficult to guarantee a robust detection based on this principle only.

More recently, Deschenes et.al. (2000) proposed a measure which looks at the change in direction between the pixel under consideration and its surrounding road pixels. The underlying hypothesis here is that at a junction, the central pixel will show a sudden change in direction, which is measured through the dot product of neighbouring eigenvectors. Fig.6b plots the direction vectors for the detected road pixels in our example. In the case of very-high-resolution images, we see a gradual change in direction much like a road curve, and we do not see the sudden change which is desirable to obtain a high detection response.

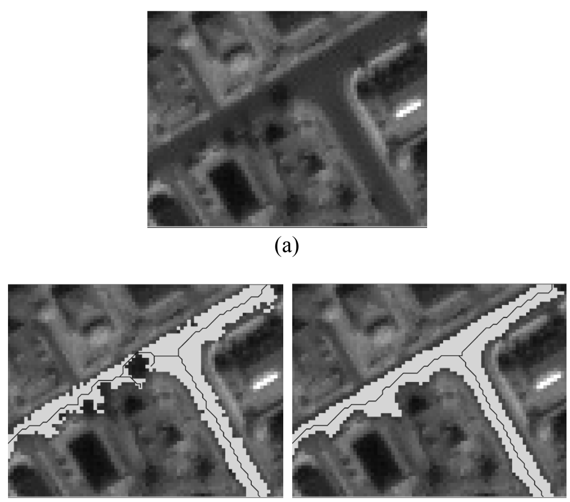
Fig.6c shows the respons of the junction detector based on improved ridge detection. In this case, we see a clear detection of the junction. In the next section, we investigate the detection response in more detail.



(a) (b)

(c)

Figure 6. Detail of the information content of a junction in very-high resolution.
(a) maximum eigenvalue, (b) direction vectors of detected ridge pixels, (c) thinned response after region growing.



(a)

(b) (c)

Figure 7. Effect of applying region growing on a smoothed image. (a) original, (b) result of region growing on original, (c) result of region growing on smoothed image

### 4.2 Junction detector using improved ridge detector

We applied our technique on an extract of a IKONOS panchromatic scene above Ghent. The extract is taken over a suburban area containing 38 road junctions (cfr Figure 2). On this extract, ridge detection was performed using a window size *w=9*. Thresholds were placed on the gradient magnitude and eigenvalues to optimize the detection rate (i.e. true positives versus false positives). Region growing was then applied to improve the result in the vicinity of junctions. We found that performing region growing on a smoothed version of the image instead of the original image improved the later steps. Especially the thinning procedure which is necessary to produce a vectorized result can be sensitive to noise in the detected boundaries of the road.

Figure 7 shows an example of the regularizing effect of (gaussian) smoothing on the vector result. Figure 7a shows the junction in the original image. Figure 7b the result of region growing and the thinned pixel chains using the intensity values of the original image. Figure 7c shows the result using the intensity values of a smoothed image. The latter shows smoother road boundaries which leads to lesser artifacts in the vector result.

Region growing is applied with an adaptive threshold *T=7*. Figure 5 shows a typical example of the detection result. To improve the robustness of the detection, we can filter out the junctions which belong to road segments of a certain minimal length. Figure 8 plots the detection rate in function of the minimum segment length. For a minimum length of 20 pixels, a true positive rate of 70% and a false positive rate of 12% is achieved. Increasing the threshold lowers the false positive rate below 10%.

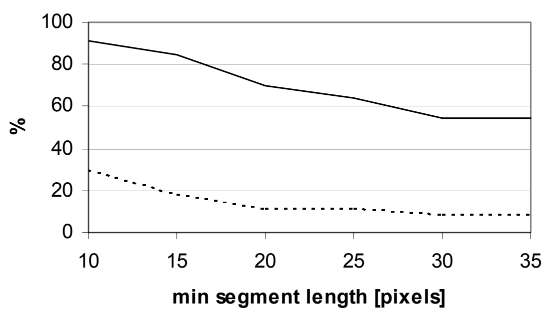



Figure 8. Detection rate versus minimum segment length. Full curve gives true positives. Dashed curve gives false positives.

## 5. DISCUSSION

In this paper, we presented a junction detector based on an improved ridge detector using region growing. The detector is specifically tuned towards detection of road junctions in very high resolution images, where we observe a clear deviation of the simple line model especially in the vicinity of junctions. Region growing offers a simple model to extend the performance of the basic ridge detector. Of special interest is the possibility to include spectral information in the detection. Whereas ridge detection is a purely geometric detector related to image structures (as is the case for e.g. a gradient operator), region growing allows us to include spectral information within the road model. In this respect, a critical parameter which has not yet been discussed in this paper is the adaptive threshold used by region growing. Within our experiments, this threshold has been set manually. In further work, the link with the spectral properties of the road class will be investigated through the use of supervised classification.
The experiments as presented in this paper show a reasonable detection of junctions. More importantly, the number of false alarms can be kept low. This is essential in our work towards the use of image derived information for automated quality assessment of GIS data. The graph matching technique that is applied to correspond the image data with the vector data is able to detect a certain amount of false alarms, but keeping this number low is essential for a reliable result.

## ACKNOWLEDGEMENTS

The authors wish to thank the support of the Federal Science Policy. This work is performed under the project "The GIS problem detector" which is part of the STEREO research programme. Special acknowledgement goes to Tele Atlas for the use of their dataset.

## REFERENCES

Arcelli C., Cordella L. P., Levialdi S. (1981) From local maxima to connected skeletons. IEEE Trans. Pattern Anal. Mach. Intell. 3(3):134-143.

Beaudet P. (1978) Rotationally invariant image operators. In: Proc. Internat. Joint Conf. Pattern Recognition, pp.579-583.

Deschenes F., Ziou D. (2000) Detection of line junctions and line terminations using curvilinear features. Pattern Recognition Letters 21:637-649.

Gautama S., Borghgraef A. (2003) Detecting change in road networks using continuous relaxation labeling. In: Proc. ISPRS Workshop "High Resolution Mapping from Space 2003". Hannover. 6 pag.

Haralick R.M., Watson L. (1981). A facet model for image data. Computer Graphics Image Proccesing, 15, pp.113-129.

Levine M., Shaheen S. (1981) A modular computer vision system for image segmentations, IEEE Trans. Pattern Anal. Mach. Intell. 3(5):540-554.

Steger C. (1998). An Unbiased Detector of Curvilinear Structures. IEEE Trans. Pattern Anal. Mach. Intell. 20(2): 113-125.

Wiedemann C. (2002) Improvement of Road Crossing Extraction and External Evaluation of the Extraction Results. In: Proc. ISPRS Symposium "Photogrammetric Computer Vision" (PCV'02), Graz 2002, Intl. Arch. Photogrammetry and Remote Sensing, Vol. 34, Part 3-B, pp. 297-300.