

# Detection and Vectorization of Roads from Lidar Data

Simon Clode, Franz Rottensteiner, Peter Kootsookos, and Emanuel Zelniker

## Abstract

*A method for the automatic detection and vectorization of roads from lidar data is presented. To extract roads from a lidar point cloud, a hierarchical classification technique is used to classify the lidar points progressively into road and non-road points. During the classification process, both intensity and height values are initially used. Due to the homogeneous and consistent nature of roads, a local point density is introduced to finalize the classification. The resultant binary classification is then vectorized by convolving a complex-valued disk named the Phase Coded Disk (PCD) with the image to provide three separate pieces of information about the road. The centerline and width of the road are obtained from the resultant magnitude image while the direction is determined from the corresponding phase image, thus completing the vectorized road model. All algorithms used are described and applied to two urban test sites. Completeness values of 0.88 and 0.79 and correctness values of 0.67 and 0.80 were achieved for the classification phase of the process. The vectorization of the classified results yielded RMS values of 1.56 m and 1.66 m, completeness values of 0.84 and 0.81 and correctness values of 0.75 and 0.80 for two different data sets.*

## Introduction

### Motivation and Goals

Road extraction from remotely sensed data is a challenging research topic that has been approached in many different ways. Compared with other remote sensing data sources, extraction of roads from light detection and ranging (lidar) data is in its infancy. Lidar sensor technology is evolving rapidly and now allows the acquisition of very dense point clouds in a short period of time (Kraus, 2002). The recent popularity of Geographic Information Systems (GIS) has stimulated research on automated road extraction in order to simplify the data acquisition and update process (Hinz and Baumgartner, 2003). Existing road extraction techniques typically produce poor detection rates and often require

existing data and/or user interaction in semi-automatic techniques (Zhang, 2003; Hatger and Brenner, 2003).

The primary goal of this paper is to present a road extraction technique that provides results of an acceptable quality but relies solely on the acquired lidar data. As lidar is both an active sensor and an explicit 3D data source, several benefits can be realised from a lidar-only approach. Data acquisition will not be limited to daylight hours as with other sensors, accurate height information is contained in the data, registration of different data sources is not required and as only one data source is used, acquisition costs are minimized.

The road extraction process will be split into two stages: classification and vectorization. Classification is accomplished by applying a hierarchical method yielding a binary image of ground elements (hence called "pixels") classified as belonging to a road. Vectorization of roads is then performed by convolving this binary image with a Phase-Coded-Disk (PCD) in order to extract the road centerline and to determine the road width. The PCD method was developed in order to overcome some of the problems encountered with existing methods. For example, the Hough transform (Duda and Hart, 1972) does not find the road centerlines in high-resolution images accurately because the peak of the Hough transform will correspond to the longest line in the image. Unfortunately, in high-resolution imagery this is the diagonal of a road segment and not the road centerline (Clode *et al.*, 2004b). Although the major emphasis of the second stage is on the vectorization of roads, the vectorization technique in this paper can be used on any binary image or irregularly spaced points to extract the centerline and width of thin elongated (but 2D) objects. The use of the PCD is not limited to extracting roads from only lidar data. It can be used for thick line parameter determination in any binary image. Examples are presented for two test sites in Fairfield (New South Wales) and Yeronga (Queensland) both covering an area of 2 km × 2 km.

### Background

The problem of road extraction from optical or radar imagery is a well-studied one which is well-summarized by Auclair-Fortier *et al.* (2000), Mena (2003), and Zhang (2003). There are many difficulties associated with detecting roads from an aerial image. The detection of the road centerline, width, and direction are all important information that is required for parameterization of the detected road. The extraction of

---

Simon Clode is with the IRIS Group, EMI, University of Queensland, Brisbane, QLD 4072, Australia (sclode@itee.uq.edu.au).

Franz Rottensteiner is at the Cooperative Research Centre for Spatial Information at the University of Melbourne, Australia (franzr@unimelb.edu.au).

Peter Kootsookos is at the United Technologies Research Center, Silver Lane, USA (KootsoP@UTRCCCT.res.utc.com).

Emanuel Zelniker is with the IRIS Group, EMI, University of Queensland, Brisbane, QLD 4072, Australia (zelniker@itee.uq.edu.au).

---

Photogrammetric Engineering & Remote Sensing  
Vol. 73, No. 5, May 2007, pp. 517–535.

0099-1112/07/7305-0517/\$3.00/0  
© 2007 American Society for Photogrammetry  
and Remote Sensing

road information from lidar data on a broad scale is still in its infancy, but existing methods suggest that the basic task is similar to other remotely sensed data. The task of the extraction of road information from remotely sensed data involves two distinct steps. An initial classification of the road network must first be achieved before the second step of vectorization can be performed. For completeness, a background of road detection and extraction will be given for a variety of sensors, including lidar, detailing the various techniques used.

An important issue in classifying techniques for road extraction is the road model and the representation of a road that is used. The selection of a road model is dependent on the appearance of a road in the sensor data, which in turn largely depends on the sensor resolution (Zhang, 2003). In low-resolution data, roads appear as relatively thin lines, and the problem of road detection becomes identical to the problem of line detection, eventually exploiting connectivity of roads in a road network for separating lines corresponding to roads from other lines. Such a road model will only deliver the centerline of a road. In high-resolution data, roads appear as two-dimensional areas rather than one-dimensional lines; they have width as well as length. The centerline can no longer be observed directly, but has to be derived by other methods, for instance from the road boundaries. Methods for road detection from high-resolution data may concentrate on detecting and grouping image edges in order to find road edges or on the spectral properties of the road surface. Whereas the accuracy of the road parameters that can be extracted from high-resolution data is superior to the accuracy achievable from low-resolution data, road detection is made more complicated by the occurrence of image edges that do not correspond to roads, but to other objects such as buildings or cars. That is why there have been efforts to combine methods using both high- and low-resolution data in an approach making use of the different appearances of roads in scale-space (Heipke *et al.*, 1997). Auclair-Fortier *et al.* (2000) divide road characteristics into four different types: spectral, geometric, topologic, and contextual, while describing the characteristics of many different detection techniques. From this summary, we can categorize road detection techniques by the road model that is applied, related to the resolution of the sensors involved as described above. We distinguish techniques modeling a road by its centerline only (typically using low-resolution imagery) from techniques delivering a more elaborate road model, where the output can be road segments, road polygons, and/or road attributes (typically using high-resolution imagery).

Wiedemann and Hinz (1999) propose to use multi-spectral satellite imagery to automatically extract roads. The satellite imagery is of varying resolution (6 m to 18 m ground pixel size), which allow the detection of the centerlines only. The local region and global properties of a road network are used to create an initial road model. The method identifies lines in each channel before fusion of individual channels is performed. The best path is derived from a graph network to provide a network of centerlines that ultimately represents the road network itself.

Road extraction in complex urban scenes was performed by Hinz and Baumgartner (2003) from multi-view aerial images with a high ground resolution. They use a road model exploiting knowledge about the radiometric, geometric, and topological characteristics of roads, making use not only of the image data, but also of a digital surface model (DSM). First, separate lanes are extracted as 2D segments, and these lanes are merged in a fusion process that makes use of the DSM. Thus, *road segments* are extracted. In an iterative way, a road network is constructed from these road segments with a relatively high success rate. The main

problems identified were the influence of large vehicles on the extraction process and the weakness of the model at intersections, which affected the linking of lane segments. Hinz (2004) uses explicitly formulated scale-dependent models to integrate detailed knowledge about roads with high-resolution aerial imagery. Additional global and local context models are defined, defining relationships between roads and background objects. Knowledge is inherited by the extraction strategy employed which defines when and how road and context models are used. Excellent results were obtained indicating that the method extracts roads even in complex environments. However, the authors acknowledge that the results were only achieved due to the expertise of the system developers in setting parameters correctly.

There are many methods that make use of both high- and low-resolution data, and varying the resolution of an image in scale space is another common approach in order to achieve this. Lee *et al.* (2000) extracted roads from 1 m resolution simulated satellite images by initially varying the scale space and applying a watershed algorithm. Grey levels and shape cues formed a basis of knowledge extraction from which the road network was modeled. It was noted that roads were not “line-like” objects in this imagery. The ultimate goal of this method was recognized to be road centerline and edge extraction, however only the identification of *road segments* was achieved.

Heipke *et al.* (1997) evaluate three different road extraction techniques. The LINE algorithm extraction is based on differential geometry, the TUM-G algorithm is based on the extraction of lines in an aerial image of reduced resolution using the approach of Steger (1996), and the extraction of edges in a higher resolution image. The TUM-S algorithm is similar to the TUM-G extraction, except that it uses ribbon-snakes to verify the roads. In the LINE algorithm, roads are represented as lines from the original 3.6 m ground pixel size image. The *completeness* rate of the LINE technique is considered high, but the *correctness* and *quality* values are considered low due to weaknesses in the road model. Many other linear structures in the image were classified as roads. Due mainly to an improved road model and a higher resolution image, the TUM-G algorithm produces significantly larger *correctness* values. As this method initially relies on line detection in low-resolution images, a reduced *completeness* value is observed as the model criteria do not match in both high- and low-resolutions. The result from the TUM-G algorithm is quadrilaterals that represent road segments, constructed from the road edges. The TUM-S algorithm also utilized the LINE algorithm in a pre-processing step. Significantly improved *completeness* and *correctness* values were obtained by discriminating roads from other “line-like” objects by considering a constant road width. Results were described also by quadrilaterals representing road segments.

Huber and Lang (2001) performed road extraction from high-resolution (2 m pixel size) airborne X-Band SAR data. Areas of interest were first identified and a road was then modeled by three homogenous regions, one central region surrounded by two adjacent regions on either side of the road by way of operator fusion. The result is complemented by using Active Contour Models to extract the final road network which is displayed as extracted road segments. The work was of a very preliminary nature, and the authors concluded that several improvements to the algorithm needed to be investigated. Priestnall *et al.* (2004) developed a framework for the extraction and classification of linear networks. The method is specifically applied to the detection of roads from panchromatic and multi-spectral imagery of spatial resolution ranging from 1 m to 30 m. The interesting concept obtained from this paper is that linear extraction generally has two steps,

identification and extraction. Extraction was performed before identification in this paper which is in contrast to the methodology employed in this paper.

There have been relatively few attempts to extract roads from lidar data. Most methods require a form of data fusion to complete the task. Estimation of road geometry parameters was performed by Hatger and Brenner (2003) in high-resolution lidar data (4 points per m<sup>2</sup>). Lidar data is used in conjunction with existing database information to derive properties such as height, slope, curvature, and width with a view to use this information in future driver information and warning systems. Although the detected properties are geometric properties of a road, they can only be detected once the lower level geometric properties such as the centerline or road segment have been determined. In this method, this information was not extracted, it was provided from an existing database. Roads were extracted from high-resolution lidar data in forested areas by Rieger *et al.* (1999). A high-quality digital terrain model (DTM) was produced by first detecting roads from which breaklines could be generated to ultimately enhance the DTM. A combination of line and point feature extraction was then used to extract the final breaklines. In their work, Rieger *et al.* (1999) used the concept of "twin snakes" to model roads as pairs of parallel edges corresponding to the road edges. Along similar lines, Akel *et al.* (2003) discuss a method for extracting a DTM in urban areas by initially estimating the DTM from the road network present. No road extraction is actually performed in this work, but the interesting assumption is made that roads lie on the DTM.

Zhu *et al.* (2004) combine high-resolution digital images and laser scan data to automatically extract roads in an urban scene. Road extraction is primarily performed from the digital images, but laser data is used to assist the extraction process by means of exploiting the explicit 3D nature of the laser data. High objects are identified within the scene and removed so that hidden road edges can be recovered. Hu *et al.* (2004) use high-resolution imagery (0.5 m pixel size) and lidar data (1.1 points per m<sup>2</sup>) for automatic urban road extraction. Road detection is primarily performed from the lidar data using both the intensity and height information. The true color imagery is used to separate grassland and trees from open areas. The model assumption that the urban roads exist in a grid structure assists the extraction of the candidate road strips that are detected by an iterative Hough transform algorithm. The work demonstrates the potential of lidar data to extract information from complicated scenes but is limited to the extraction of grid roads.

Alharthy and Bethel (2003) present a simple and fast method to detect roads in urban areas from lidar data. The main aim of the work was to exclusively use lidar data so that limitations of availability of other sources, such as ground plans, could be avoided. Both the intensity and height information were used to filter the raw lidar data and remove "noise" that was unrelated to the road class. Filtering is performed by locating penetrable objects from the first and last pulse observations. These objects are considered to be non-road regions, however, this model assumption ignores the existence of overhead powerlines and overhanging trees. This is an occurrence which happens regularly in an urban road network and should not be ignored. Refinements to the heuristic classification were then used to improve the overall classification. The results of the classification are visually displayed, but no formal analysis of the quality of the classification is carried out. Clode *et al.* (2004a) perform road classification in a manner similar to Alharthy and Bethel (2003) in that only high-resolution lidar data is used in the classification scheme. Again, both intensity and height information are used in the

classification, but the idea of a local point density is introduced. The local point density is an indicator of how many neighboring lidar points have similar spectral and geometric properties to the lidar point in question. The fact that roads are consistent in nature is an important model assumption.

Road centerline extraction is a common goal for extracting road parameter information which is not a straightforward task. Common line detection methods such as the Hough transform (Duda and Hart, 1972) detect the longest one pixel thin line in a road segment rather than the centerline (Clode *et al.*, 2004b). In low-resolution images the centerline and the longest line are the same, however in high-resolution images, they are not. In order to use a Hough transform successfully for road parameter extraction, a change in scale space must be applied but other problems are encountered. The use of lower resolution images degrades the accuracy of any detection methods used, and thus generally requires another higher resolution image to be used simultaneously during the extraction process (Heipke *et al.*, 1997). Other road parameters, such as the width and direction, are important information that can also be extracted from a high-resolution image, although many of the methods described so far do not extract some or all of this information directly, (Auclair-Fortier *et al.*, 2000). Hu and Tao (2003) extract main-road centerlines from high-resolution satellite imagery (1 to 4 metres). Linear features are sought globally in the image and grouping is performed hierarchically integrating multiple clues into the algorithm.

Standard image processing techniques have seen a range of objects detected in images. Circles were detected in digital images by Atherton and Kerbyson (1999) by introducing a complex phase coding along "spokes" in an annulus operator. The size of the circle being detected was represented by the phase coding along the spokes. Convolution using a kernel similar to that described will result in a peak position at the center of the circle and the phase at the peak representing the detected circle size. By combining the ideas of Atherton and Kerbyson (1999) with those of Huber and Lang (2001), the development of the Phase Coded Disk (PCD) was inspired for vectorization of roads (Clode *et al.*, 2004b). The PCD is a complex kernel which uses phase to code for the angle of the line. By convolving the original image with a PCD, the centerline, direction, and width can be accurately extracted at any point along the detected centerline.

The quality of documented road extraction techniques is something that has been poorly quantified in many existing methods, with a few notable exceptions. Hinz and Wiedemann (2004) recognize that automatic object extraction is not expected to deliver perfect results. Methods of internal evaluation are tested on a series of aerial images and demonstrate how confidence intervals can increase system efficiency and lessen the burden placed on human operators by highlighting problem areas and areas of uncertainty. The results of the self-diagnosis are evaluated against a manually-plotted reference with encouraging results. Wiedemann (2003) describes a two-step process that is required for evaluation of road extraction results. Both extracted and reference data are compared in vector format where the first step is to match the two data sets while the second is to calculate the quality measures. The evaluation methodology was tested on both simulated and real extraction results, and it was concluded that the evaluation scheme adequately captures the characteristics of the extraction results. Consequently, the methods described in Wiedemann (2003) and Heipke *et al.* (1997) will be used to form the basis of techniques to quantify our road extraction results.

This paper is structured as follows. The next section describes our hierarchical classification technique that

progressively classifies the lidar points by using features such as intensity data, the closeness to the DTM, the lidar point density, and the connectivity of roads. Vectorization of the classified road image is discussed in the Vectorization Section, along with the creation of the PCD and the method of convolving the image with that complex kernel. Results from the Fairfield and Yeronga data sets are discussed in the Experiments Section, while conclusions and future work are examined in the final section.

## Classification of Roads from Lidar Data

The classification algorithm presented in this paper uses a road model that exploits the continuous homogeneous nature of a road by interrogating the normalized local point density of lidar points that lie on or near the DTM and meet certain reflectance requirements in the wavelength of the lidar system. A binary classified image can be generated based on whether or not each pixel fulfils these requirements. In this binary image, the roads appear as visible thick lines of “road pixels” which form a road network. The algorithm requires raw last pulse lidar data that contain both height and intensity information. The classification of roads can be considered to have two phases, the generation of a DTM and the road classification itself. These two stages will be discussed in the following sections.

### Generation of a Coarse DTM

Our road detection algorithm requires a relatively good approximation of the terrain in order to separate points on the terrain from off-terrain points. Thus, preprocessing is required in order to create a DTM. First, a DSM represented by a height grid is interpolated from the last pulse lidar points by inverse distance weighting. The DSM represents the surface from which the laser pulse is reflected and typically contains trees, buildings, and the terrain surface. For classification purposes, morphologic filtering is often used for DTM generation (Weidner and Förstner, 1995). The size of the structural element for morphologic filtering should be small to obtain a good approximation of the terrain. However, this causes problems in areas with large (e.g., industrial) buildings where a small structural element might “fit” into such a building, and thus not eliminate it. This would produce a “hill” in the DTM. If the DTM is created by morphological filtering with a single structural element size, then the size of the structural element must be at least as large as the largest building within the data set in order to adequately remove all buildings. Unfortunately, terrain structures smaller than the structural element will also be eliminated. The result is that small hills will have their tops removed and the resultant DTM will be biased dependant on the structural element size. That is why a method based on hierarchical application of morphological greyscale openings is used, with structural elements of different sizes.

The hierarchical DTM creation method commences by creating an initial coarse DTM from one large structural element that is larger than the largest building in the data set. A rule-based algorithm is then used to detect large buildings in the data (Rottensteiner *et al.*, 2003). A smaller structural element is used to create a finer DTM, but buildings detected in the previous iteration have their corresponding heights substituted from the DTM of the previous iteration. The process is continued until a minimum size for the structural element is reached. In order to remove any of the biases that may still be present from the smallest structural element size, the final DTM is created by again interpolating by inverse distance weighting, but excluding lidar points classified as “off-terrain.” Figure 1 shows the interpolated DSM, an intermediate DTM, intermediate building detection results, and the final DTM.

### Road Classification

To extract roads from a lidar point cloud, a pipeline classification technique is used to progressively classify the lidar points into the classes “road” or “non-road.” For the purpose of this paper, we will describe any lidar data point  $p_k$  by Equation 1:

$$p_k = (x_k, y_k, z_k, i_k). \quad (1)$$

The values  $x_k$ ,  $y_k$ , and  $z_k$  represent the 3D coordinates, and  $i_k$  represents the intensity of the last pulse strike. Let  $S$  represent the set of all laser points collected as described in Equation 2:

$$S = \{p_1, p_2, p_3, \dots, p_N\}. \quad (2)$$

The values  $p_1$ ,  $p_2$ ,  $p_3$ , . . . . ,  $p_N$  are the individual lidar points.

In accordance with our road model, roads lie on or near the DTM. This is true except for elevated roads, bridges, and tunnels. We can thus disregard all lidar points that lie outside a given tolerance of the DTM (Akel *et al.*, 2003) and create a subset of lidar points as defined by Equation 3:

$$S_1 = \{p_k \in S : |z_k - DTM(x_k, y_k)| < \Delta h_{max}\}. \quad (3)$$

The value  $DTM(x_k, y_k)$  is the height value of the smoothed DTM at location  $(x_k, y_k)$ , and  $\Delta h_{max}$  is the maximum allowable difference between  $z_k$  and the DTM. This condition implicitly addresses the geometric road property requirements as the generated DTM is a smoothed surface.

The pipeline process then filters lidar points based on their intensity values. Unprocessed intensity images for the two test data sets can be seen in Figure 2. The road network is clearly visible in both datasets as dark thick connected lines throughout each scene. The algorithm requires some training in order to determine the reflectance properties of the road material to be detected. Lidar points that have last pulse intensity values that appropriately represent the reflectance properties of the road material form a new subset of the data as described by Equation 4:

$$S_2 = \{p_k \in S_1 : i_{min} < i_k < i_{max}\}. \quad (4)$$

The values  $i_{min}$  and  $i_{max}$  are the minimum and maximum acceptable lidar intensities at any point  $p_k$ .

Even though the intensity values returned by the scanning unit are typically considered under-sampled and noisy (Vosselman, 2002), the effect is minimized by the typically uniform and consistent nature of road material along a section of a road. The result of Equation 4 is a set of lidar points ( $S_2$ ) that were reflected from the road along with some other false positive (non-road) detections. If more than one type of road material is to be detected in the surveyed region, different subsets  $S'_2$  can be created according to the individual reflectance properties of the different materials being detected. A combined and complete  $S_2$  can be created by taking the union of the different subsets  $S'_2$ .

Roads are depicted as a continuous network of pixels which form thick lines. Due to this continuous nature of a road network, lidar points that have struck the middle of the road are expected to be surrounded by other points that have struck the road. That means that all lidar points in the local neighborhood (e.g., defined by a circle of radius  $d$  around any point  $p_k$ ) are expected to also be classified as “road points” according to Equations 3 and 4, with the exception of noise which would include reflections from vehicles and overhanging trees. We will call the percentage of “road points” within a local neighborhood the “local point density.” For points in

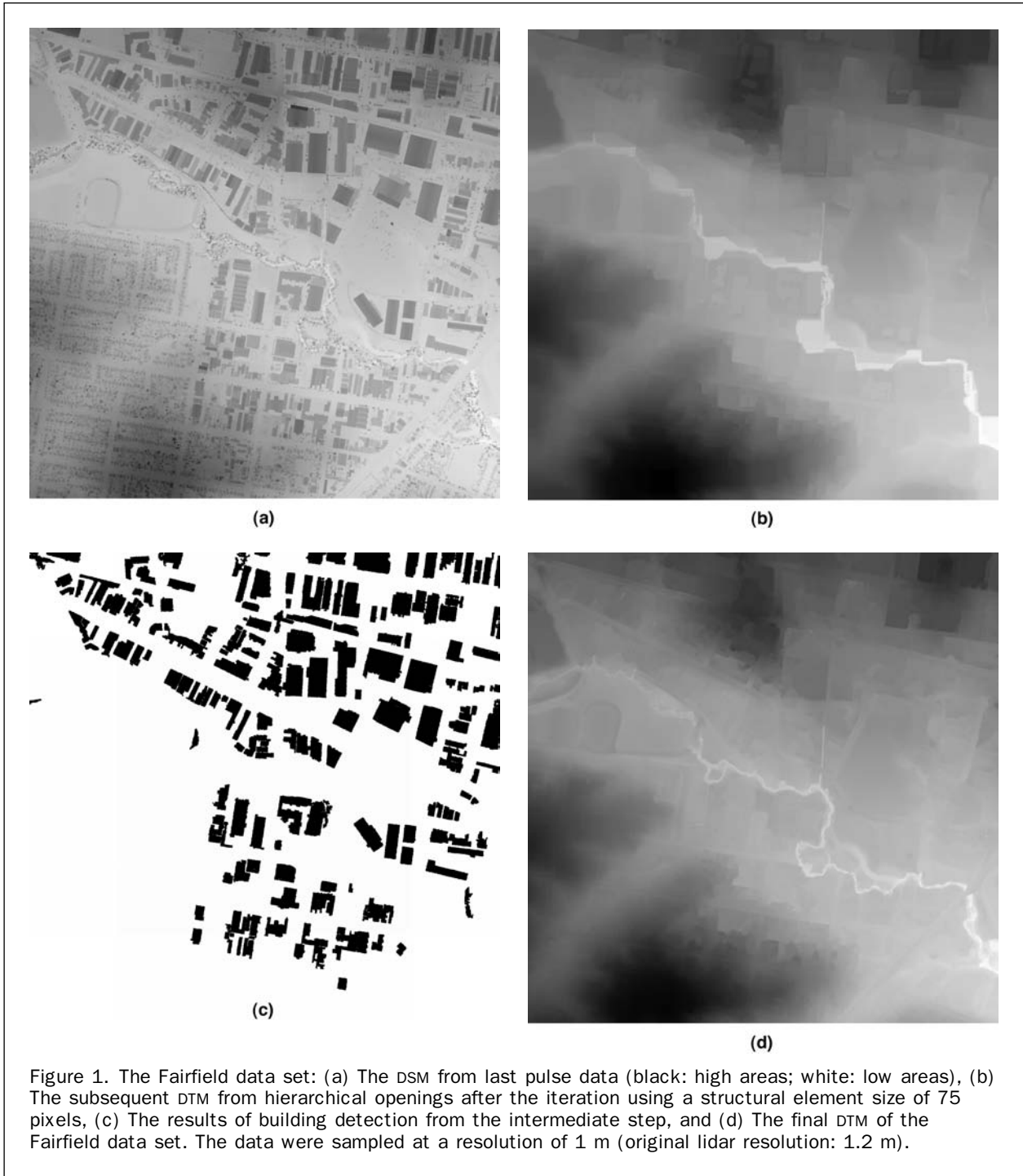


Figure 1. The Fairfield data set: (a) The DSM from last pulse data (black: high areas; white: low areas), (b) The subsequent DTM from hierarchical openings after the iteration using a structural element size of 75 pixels, (c) The results of building detection from the intermediate step, and (d) The final DTM of the Fairfield data set. The data were sampled at a resolution of 1 m (original lidar resolution: 1.2 m).

the middle of a road, the local point density should be close to 100 percent. For any point that lies on the edge of a road it can be considered to be close to 50 percent, and 25 percent for a lidar point in the corner of a sharp 90° bend. Figure 3 describes the local point density for these three cases.

Typically, we would expect between 50 percent and 100 percent of the circle to lie on the road. By testing all points against a chosen minimum local point density  $\rho_{min}$ , a new subset of points,  $S_3$ , is described as per Equation 5:

$$S_3 = \{p_k \in S_2 : \frac{|\{p_j \in S_2 : \|p_k - p_j\|_2 < d\}|}{|\{p_j \in S : \|p_k - p_j\|_2 < d\}|} > \rho_{min}\}. \quad (5)$$

The value  $d$  is the maximum distance from  $p_k$  or the radius of the local neighborhood,  $|\{ \dots \}|$  denotes the number of points  $p_j$  in the respective set, and  $\|p_k - p_j\|_2$  is the Euclidean distance from  $p_j$  to  $p_k$ . An upper bound should be placed on the possible values of  $d$  so that it is any value less than or equal to half of the expected maximum road width.

A binary image,  $F(x, y)$ , is now created from the final subset  $S_3$  with a pixel size  $\Delta$  loosely corresponding to the original average lidar point density. The pixel values  $f(x, y)$  of the binary image are determined according to whether a point  $p_k \in S_3$  exists inside the area represented by the pixel at position  $(x, y)$  or not:



(a)



(b)

Figure 2. Lidar intensity images: (a) Fairfield, and (b) Yeronga.

$f(x,y) =$

$$\begin{cases} 1 & \text{if } \exists p_k \in S_3 | \left( x - \Delta/2 < x_k \leq x + \Delta/2 \right) \wedge \left( y - \Delta/2 < y_k \leq y + \Delta/2 \right) \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

where  $\wedge$  is the logical AND operation. The “white” pixels in that binary image, characterized by  $f(x, y) = 1$ , thus represent roads. Many small gaps exist between these “road pixels.” There are several reasons for this. First, the creation

of an image with a pixel size smaller than the lidar point density in some locations of the image will mean that the lidar data is under-sampled in this location. Second, reflections from other objects such as vehicles and overhanging trees will also cause gaps to appear.

These gaps are removed using a two-step approach based on morphologic filtering. First, a morphological closing with a small structural element is initially performed to connect neighboring road pixels. The Not-Road image (i.e.,  $1 - f(x,y)$ ) is then labelled using a connected component analysis, and the values of all pixels corresponding to Not-Road segments with a small area are switched to 1 in the binary image  $f(x,y)$ , which results in a “road image” that contains all public roads, private roads, car parks, and some noise. In a second stage of processing, another label image is created from this binary image in order to identify individual continuously connected road segments. This time, small road segments are erased in the binary image of road pixels, thus removing most of the noise present. This ensures that our detected roads meet the requirement of our road model, i.e., to be continuous in nature.

Depending on the definition of the road model, car parks can be considered to be either included in the road model or not. Unfortunately, it is very difficult to distinguish a car park from a road as both have similar surface and reflectance properties. By defining a maximum acceptable road width prior to processing, (which is used in the next section to determine the size of the PCD) very wide unconnected car parks can be removed from the binary image. As roads form a network of long thin connected objects, the area ratio of each individual road segment and the corresponding minimum bounding rectangle (MBR) will decrease as the length of the smallest side in the MBR increases. Large isolated blobs can be detected in the image using this ratio, thus allowing the removal of any unconnected car parks from the final binary image of road pixels.

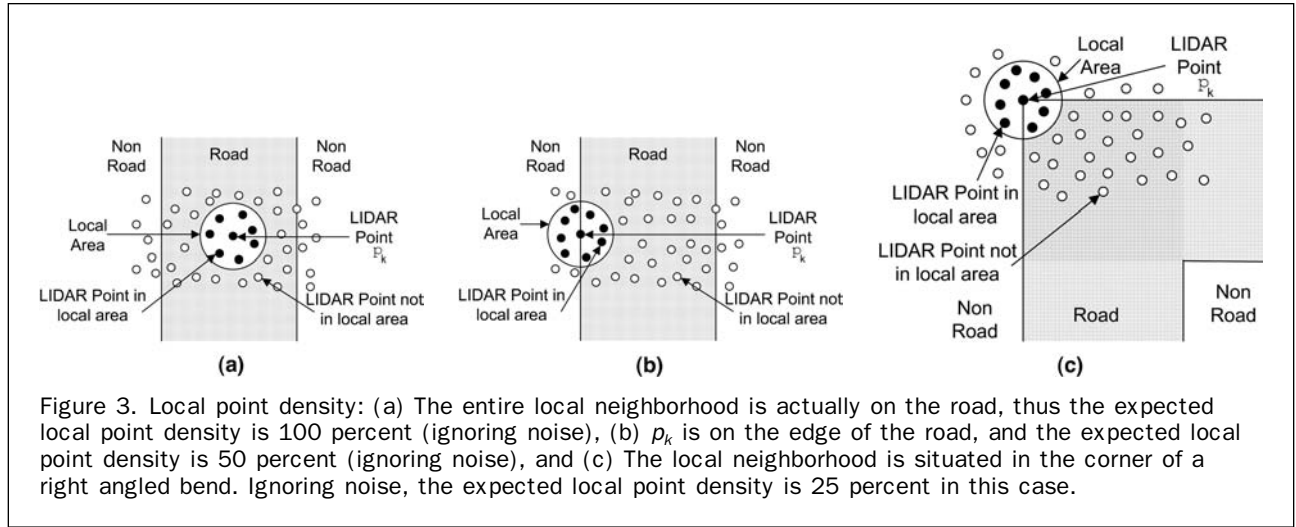
### Vectorization of Classified Road Images

Road vectorization from high-resolution imagery requires thick line parameter estimation where the parameters being sought are the centerline, orientation  $\phi$ , and width  $w$ . A method for thick line parameter extraction is presented in this section with specific application to the vectorization of the detected roads obtained from the previous section. The road parameters are determined by considering only the local neighborhood around each pixel, thus removing the influences of other road segments. Road parameters are determined within the local neighborhood using a convolution of the binary image with a Phase Coded Disk (PCD). The radius of the PCD,  $r$ , determines the size of the local area to be used in parameter detection. A tracing algorithm is then used to vectorize the detected road parameters, ultimately yielding a vectorized road network.

This section will first describe the road centerline, orientation  $\phi$ , and width  $w$ , in terms of the local neighborhood. The creation of the PCD and the convolution process is then described. The calculation of each of the road parameters from the convolution results are described along with the tracing algorithm used to ultimately vectorize the road network. For the purpose of discussion in this section, the binary image of road pixels defined according to Equation 6 is assumed to be continuously valued in spatial coordinates.

### Defining the Road Parameters

As defined by Clode *et al.* (2004b), the road centerline, orientation  $\phi$ , and width  $w$  are three critical parameters that need to be determined in order to be able to model and



ultimately vectorize the binary road image. Figure 4a illustrates the local neighborhood around a point as a disk of radius  $r$  from a point on the centerline of a road. This can be considered to be the area covered by a PCD of radius  $r$  being overlaid on the centerline of a road. The orientation of the road, denoted  $\phi$ , is the angle between the coordinate axes and the road centerline. The width of the road is described by  $w$ . Road edges can be calculated by applying half of the width orthogonally to the road orientation at the PCD location.

#### The PCD and the Convolution

This section describes the construction of the PCD and the convolution.

##### Defining the PCD

The PCD that we propose is that of Clode *et al.* (2004b) and is defined by Equation 7:

$$O_{PCD} = e^{j2 \cdot \tan^{-1}(b/a)} = e^{j2\vartheta}. \quad (7)$$

The variables  $a$  and  $b$  are  $x$  and  $y$  coordinates relative to the center of the PCD. Further,  $a^2 + b^2 \leq r^2$ ,  $\vartheta = \tan^{-1}(b/a)$ ,  $j^2 = -1$ , and  $r$  is the radius of the disk. The constant 2 in the exponent has been introduced into the definition of the PCD in order to ensure that pixels that are diametrically opposite in their direction from the center of the kernel (e.g.,  $30^\circ$  and  $210^\circ$ ) during the convolution process indicate the same direction (e.g.,  $30^\circ$ ) after convolution and do not cancel out. Figure 4b shows a constructed PCD. The shading and scale bar illustrate the phase coding. The amplitude is unity across the disk.

##### Convoluting the PCD with the Binary Road Image

The convolution of the PCD with the binary image takes the form:

$$Q(x, y) = F(x, y) \otimes O_{PCD} \quad (8)$$

where  $Q(x, y)$  is the resultant image,  $\otimes$  is the convolution operation,  $O_{PCD}$  is the PCD, and  $F(x, y)$  is the binary road image. As described in Clode *et al.* (2004b), the convolution is rotationally invariant with respect to the orientation of the road, which allows the extraction of the line at any orientation. The result of the convolution defined in Equation 8 yields a magnitude and a phase image that are defined by Equations 9 and 10, respectively:

$$M = |F(x, y) \otimes O_{PCD}|, \text{ and} \quad (9)$$

$$\phi = \frac{1}{2} \arg(F(x, y) \otimes O_{PCD}). \quad (10)$$

From the result of the convolution, the magnitude (Equation 9) and phase (Equation 10) images can be used to determine the three desired parameters of a road as defined by Clode *et al.* (2004b). The result of the convolution can be indirectly considered to represent the road parameters we desire. The relationships between the magnitude and phase images and the road parameters are detailed as follows.

##### Determining the Road Parameters by Convolution with the PCD

This section describes the extraction of each of the three road parameters and the convolution result. Before the appropriate relationships can be obtained, the complex integral of the entire PCD must be evaluated only over the area of the PCD that coincides with road pixels. The integral is first modified before the three road parameters are discussed.

##### Defining the Integral Over Only the Road

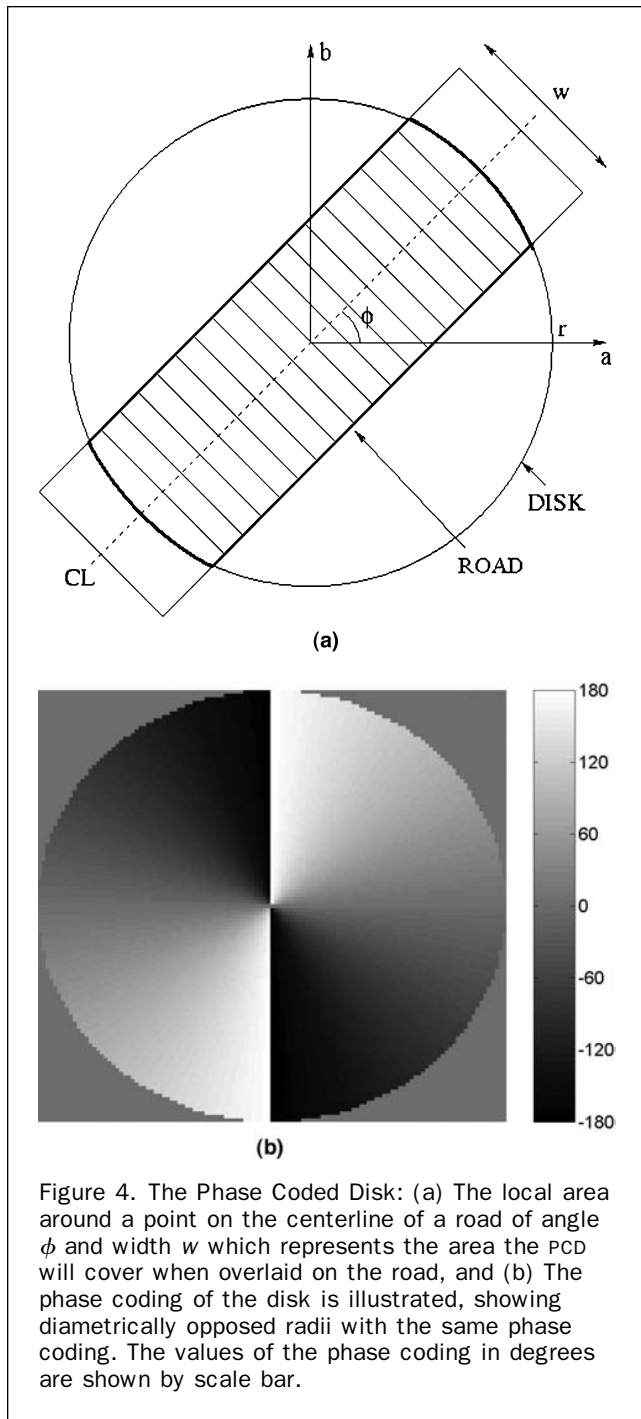
The complex integral over the entire disk at any position  $(x, y)$  is described by Equation 11:

$$q(x, y) = \int_{-\pi}^{\pi} \int_0^r f(x, y) \cdot e^{j2\vartheta} \cdot u \cdot du \cdot d\vartheta. \quad (11)$$

The variable  $u$  is a substitute variable that has been introduced to represent the radius of the PCD as  $r$  is in the limits of the integral and the function  $f(x, y)$  is understood to be translated to appropriate polar coordinates.

In order to evaluate the integral over only the road area covered by the disk (i.e., all areas where  $f(x, y) = 1$  as described in Equation 6, the limits of the inner most integral are changed so that the area defined is only the road contained within the disk as opposed to the entire disk as described in Equation 11. The described area is illustrated in Figure 5a. It can be seen that there are two distinct cases for areas of road covered by the PCD. These two distinct cases are shaded differently in Figure 5a. Areas of the road shaded lightly can be integrated over the full radius of the PCD, however, the dark shading represents areas that can only be integrated to a distance of  $\frac{w}{2 \cos \vartheta}$ . The variable  $R$  is used to represent the limits of



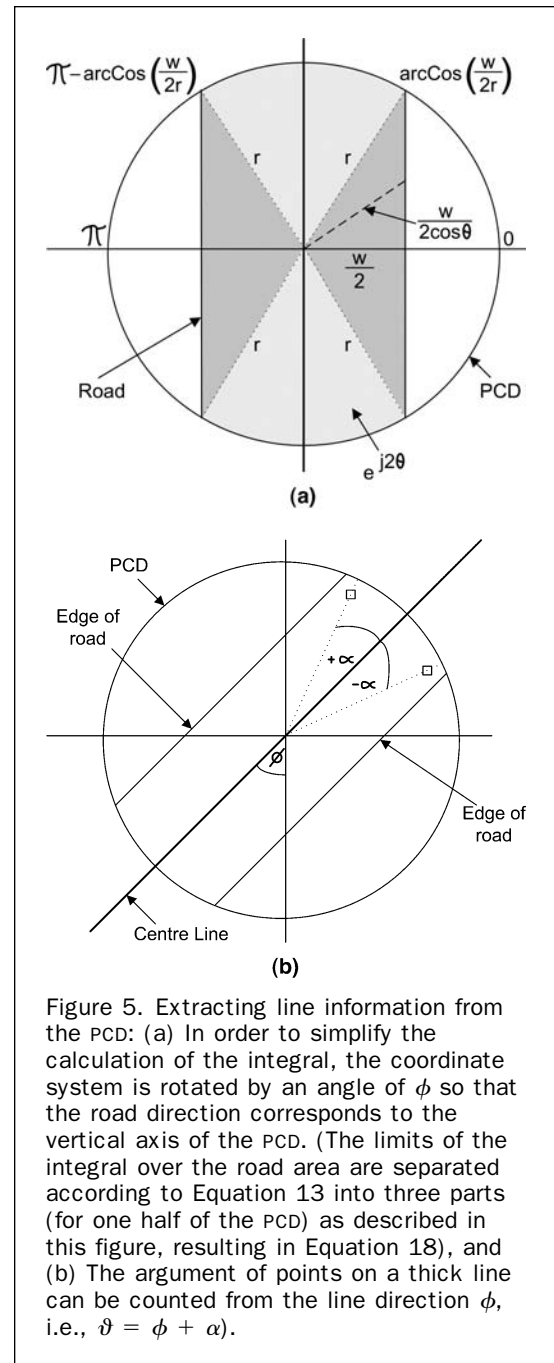


the integral at all different angles of  $\vartheta$  over the road section contained within the PCD and is described by Equation 12:

$$R = \min\left(\frac{w}{2\cos\vartheta} \cdot r\right). \quad (12)$$

Thus, the integral over the road can be described as the integral over the disk where  $f(x, y) = 1$  as described in Equation 13:

$$q(x, y) = \int_{-\pi}^{\pi} \int_0^{R_r} f(x, y) \cdot e^{j2\vartheta} \cdot u \cdot du \cdot d\vartheta. \quad (13)$$



#### Centerline

The magnitude (Equation 9) of the convolution will be at a maximum when the PCD is centered on the road centerline and the area of the road within the PCD is half of the area (as opposed to half of the diameter, i.e., the radius) of the PCD itself. This is described by Figure 5 where a PCD is centered on a road centerline. The magnitude image is not only a centerline indicator but also a centerline certainty indicator. Obviously, the magnitude value will be scaled by the size of the defined PCD.

The magnitude of the convolution quantifies the amount of support there is for the existence of a road at the determined road orientation. Essentially, the component of each pixel along the determined line orientation is compared with the component of each pixel across the determined line orientation, thus providing an overall support for this road



direction. Equation 16 has a period of  $\pi$  or  $180^\circ$ , thus pixels that are different by  $180^\circ$  will accumulate, and pixels that are  $90^\circ$  different will cancel. The magnitude will be a maximum when the majority of road pixels under the disk contribute to that particular road orientation. A road with a width of one pixel will have a magnitude less than a slightly thicker road, as there are more pixels that can positively contribute towards the magnitude. As the line width increases, so does the magnitude but at a slower rate until the area of the road in the PCD is the same as the area of non-road in the PCD. If the road width increases past this point, then any additional pixels do not positively contribute to a line running in that direction, i.e., they contribute more to a line running at  $90^\circ$  to the road orientation ( $\phi$ ).

This fact imposes a constraint on the radius of the PCD,  $r$ , such that it is must be larger than the maximum road width to be detected.

Points off center will have a magnitude value less than that at the centerline thus creating a ridge within the magnitude image that represents the centerline itself. At junctions and intersections the resultant magnitude will be smaller or even zero due to competing influence of multiple road segments in different directions.

#### Road Orientation

Extracting the orientation of the road from the convolution is achieved by first realizing that when the PCD is centered on the centerline of the road, symmetry is maintained. The phase of  $q(x, y)$  at any position of the ridge previously discussed can be shown to be twice the line orientation by substituting  $\vartheta = \phi + \alpha$  and  $d\vartheta = d\alpha$  into Equation 13. Each road pixel will now be at an orientation  $\alpha$  to the line direction  $\phi$ . Equation 14 redefines the integral at any pixel  $q(x, y)$  in image  $Q(x, y)$  to incorporate the line orientation  $\phi$ :

$$q(x, y) = \int_{-\pi}^{\pi} \int_0^R e^{j2(\phi+\alpha)} \cdot u \cdot du \cdot d\alpha. \quad (14)$$

This simplifies by taking all elements in the equation that do not depend on  $u$  or  $\alpha$  outside of the integral to give:

$$q(x, y) = e^{j2\phi} \int_{-\pi}^{\pi} \int_0^R e^{j2\alpha} \cdot u \cdot du \cdot d\alpha. \quad (15)$$

Equation 15 can be analyzed in a purely real form by utilizing the symmetry of the PCD, demonstrated in Figures 4 and 5, along with the fact that the results of each half of the integral are complex conjugates of the other. As such, Equation 15 is rewritten in Equation 16:

$$q(x, y) = e^{j2\phi} \int_0^{\pi} \int_0^R 2\cos 2\alpha \cdot u \cdot du \cdot d\alpha. \quad (16)$$

The phase of  $q(x, y)$  in Equation 16 does not depend on the result of the integral (except if it is zero) because the integral is purely real. Thus, the phase of  $q(x, y)$  at any position represents twice the directional angle  $\phi$  of the underlying road.

#### Road Width

In order to calculate the width of the road the radius of the PCD,  $r$ , should be at least as large as the maximum line width to be detected as previously described. Figure 5 shows an example of a disk of radius  $r$  overlaid on a thick line of width  $w$  and orientation  $\phi$ . The convolution can be calculated by organizing the complex integral into a purely real integral again based on the symmetry of the PCD as

discussed in the proceeding section. The real integral over the road is described in Equation 17:

$$q(x, y) = \int_0^{\pi} \int_0^R 2\cos 2(\vartheta) \cdot u \cdot du \cdot d\vartheta. \quad (17)$$

Rewriting Equation 17 as per Figure 5 and Equation 12 yields Equation 18:

$$\begin{aligned} q(x, y) = & \int_0^{\cos^{-1}\left(\frac{w}{2r}\right)} \int_0^{\frac{w}{2\cos\vartheta}} 2\cos 2\vartheta \cdot u \cdot du \cdot d\vartheta \\ & + \int_{\cos^{-1}\left(\frac{w}{2r}\right)}^{\pi - \cos^{-1}\left(\frac{w}{2r}\right)} \int_0^r 2\cos 2\vartheta \cdot u \cdot du \cdot d\vartheta \\ & + \int_{\pi - \cos^{-1}\left(\frac{w}{2r}\right)}^{\pi} \int_0^{\frac{w}{2\cos\vartheta}} 2\cos 2\vartheta \cdot u \cdot du \cdot d\vartheta. \end{aligned} \quad (18)$$

By evaluating all terms, Equation 18 evaluates to Equation 19:

$$\begin{aligned} q(x, y) = & -\frac{w^2 \sin\left(\cos^{-1}\left(\frac{w}{2r}\right)\right)}{2} \frac{2r}{w} \\ & + w^2 \cos^{-1}\left(\frac{w}{2r}\right) - 2r^2 \sin\left(\cos^{-1}\left(\frac{w}{2r}\right)\right) \frac{w}{2r}. \end{aligned} \quad (19)$$

Equation 19 can be shown to simplify to Equation 20:

$$q(x, y) = w^2 \cos^{-1}\left(\frac{w}{2r}\right) - 2w \sqrt{r^2 - \frac{w^2}{4}}. \quad (20)$$

As this convolution is purely real, i.e., no imaginary parts, the absolute value of this integral is the same as the magnitude of the convolution, thus the magnitude of the convolution is described in Equation 21 in terms of the width of the road  $w$  and the radius of the PCD  $r$ :

$$M = \left| w^2 \cos^{-1}\left(\frac{w}{2r}\right) - 2w \sqrt{r^2 - \frac{w^2}{4}} \right| \quad (21)$$

The width of the line can be calculated from the relationship between the magnitude  $M$ , the width of the road  $w$ , and the radius of the PCD,  $r$ . A graph can be generated for a PCD of a fixed radius thus enabling the width of a road to be determined at any point. Figure 6 illustrates the relationship between width and magnitude and that a level of ambiguity resolution is required (e.g., a magnitude of 600 could imply a width of 17 or 32). A constraint placed on  $r$  will resolve this ambiguity as well as avoid saturation problems that will occur when the road width is greater than the kernel radius.

#### Vectorizing the Results of the Convolution

The final road network consists of a set of road segments, each of them represented by three polylines, namely the centerline and two road edges. Vectorization consists of

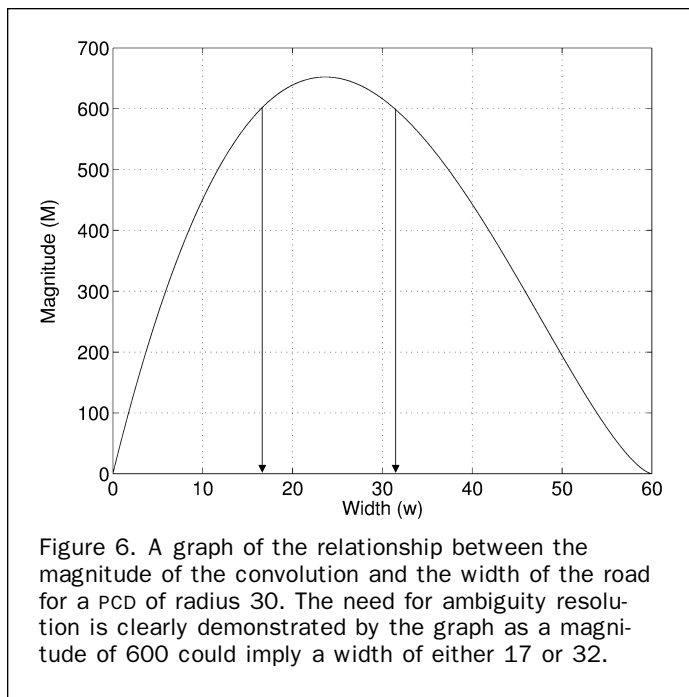


Figure 6. A graph of the relationship between the magnitude of the convolution and the width of the road for a PCD of radius 30. The need for ambiguity resolution is clearly demonstrated by the graph as a magnitude of 600 could imply a width of either 17 or 32.

three steps. First, the road centerline of each road segment is extracted. Second, the extracted road centerlines are joined or intersected with other centerlines of the neighboring segments to create a continuous network of road centerlines. An exact intersection is created by splitting the crossing road centerlines at the intersection point. Where a T-Junction is encountered, the road segment at the base of the T is extended straight ahead until the centerlines of both roads intersect. A list of all crossing points and T-Junctions is kept in conjunction with the connecting road segments in order to maintain the topology of the road network. Third, the polylines representing the road edges are constructed as being parallel to the centerline at a distance of half the road width. The polygon formed by the two edge polylines defines the road segment. The set of all road segments and their topology describes the continuous road network.

#### Vectorizing the Road Centerlines

The ridge of the magnitude image (as described in the Centerline Section) is traced in order to extract the road centerlines. Tracing is achieved by initially masking the magnitude image with the binary road mask in order to limit the search space and remove noisy edge areas. The maximum of the magnitude image is found and the corresponding line direction is read from the phase image. The tracing algorithm moves along the line pixel by pixel ensuring that the point is still a maximum against its neighbors until the line ends at a pixel of zero magnitude. A polygon is created that consists of a series of centerline segments. Points to the side of each centerline segment within the calculated road width are zeroed as the centerline is traced. The zeroing indicates that this portion of the road has been completed and ensures a similar path is not retraced. Once the road segment is completed the process is repeated from the original maximum but with the diametrically opposed line direction. Then, the maximum of the remaining untraced magnitude pixels is found and tracing is recommenced from this pixel. The process is repeated until all relevant pixels have been traced.

Thus, one centerline segment after the other is extracted until the masked magnitude image is completely blank. This ensures that all ridges have been traced. The next important

step is to connect the neighboring road segments. Due to noise in the magnitude image, the tracing of some road centerlines is terminated prematurely by the tracing algorithm. In such cases, the tracing algorithm will extract another centerline that will terminate close to the original prematurely terminated centerline end. Connection of road centerline ends is performed by concatenating road segment chains that have ends that are both close and pointing to each other as described in Figure 7a. The segments are close if they are within one road width of the centerline end. A check is made that the midpoint of the link actually lies on a ridge in the original magnitude image ensuring that dead end streets do not erroneously get connected as described in Figure 7b. Once two road segments have been connected the process is repeated until no more concatenation can occur. During concatenation, all road crossings are found by identifying crossing road segments. These segments are split at the intersection and the road crossing is created (Figure 7c), leaving only the determination of T-junctions. This requires that the end of each centerline be checked against all other centerlines. In a manner similar to the way individual centerlines were concatenated to form longer centerlines, the closest point on any other centerline is found for each centerline end point. If the direction to the closest other centerline position matches the current centerline direction, the road segment is extended if the midpoint corresponds to road in the binary image (Figure 7d). The process is repeated until all T-junctions have been modeled.

#### Vectorizing the Road Edges

The edges of the road are also represented as polygons that are created by calculating the width of the road and the direction of the road  $\phi$ . The road width is then smoothed by applying a low pass filter to the widths of each road segment. At each point along the centerline of each road segment, two new road edge points are created based on the road direction plus or minus  $90^\circ$  ( $\phi \pm \pi/2$ ) and half the smoothed road width ( $w/2$ ). Two edge polylines are defined for each road segment thus defining the road segment polygon. At the end of each centerline, the intersections of the accompanying road edges are calculated. The edges within the intersection are kept to ensure that the connection between the centerline and the edges within a road segment are kept intact. To complete the visualization at the intersections, blanked road edge ends on the same side of the centerline are then joined.

## Experiments

#### Description of the Data Set

Test data was captured over two sites, Fairfield (NSW) and Yeronga (QLD) using an Optech ALTM 3025 laser scanner. Each site covers an area of  $2 \text{ km} \times 2 \text{ km}$ . Both first and last pulses and intensities were recorded with an average point distance of about 1.2 m and 0.5 m, respectively. The Fairfield data set changes dramatically in its nature throughout the area. Land-use changes from urban to rural-like with both residential and industrial urban areas present. The industrial regions have larger buildings with many car parks and private roads while the residential regions have a much smaller average building and block sizes. The Yeronga data set changes less dramatically, but is an urban area with both industrial and residential use.

A reference data set was created for both Fairfield and Yeronga by digitizing roads interactively from a digital orthophoto of the area that had been created with a resolution of 0.15 m. The guideline used during digitizing was that public roads were to be classified as roads, but car

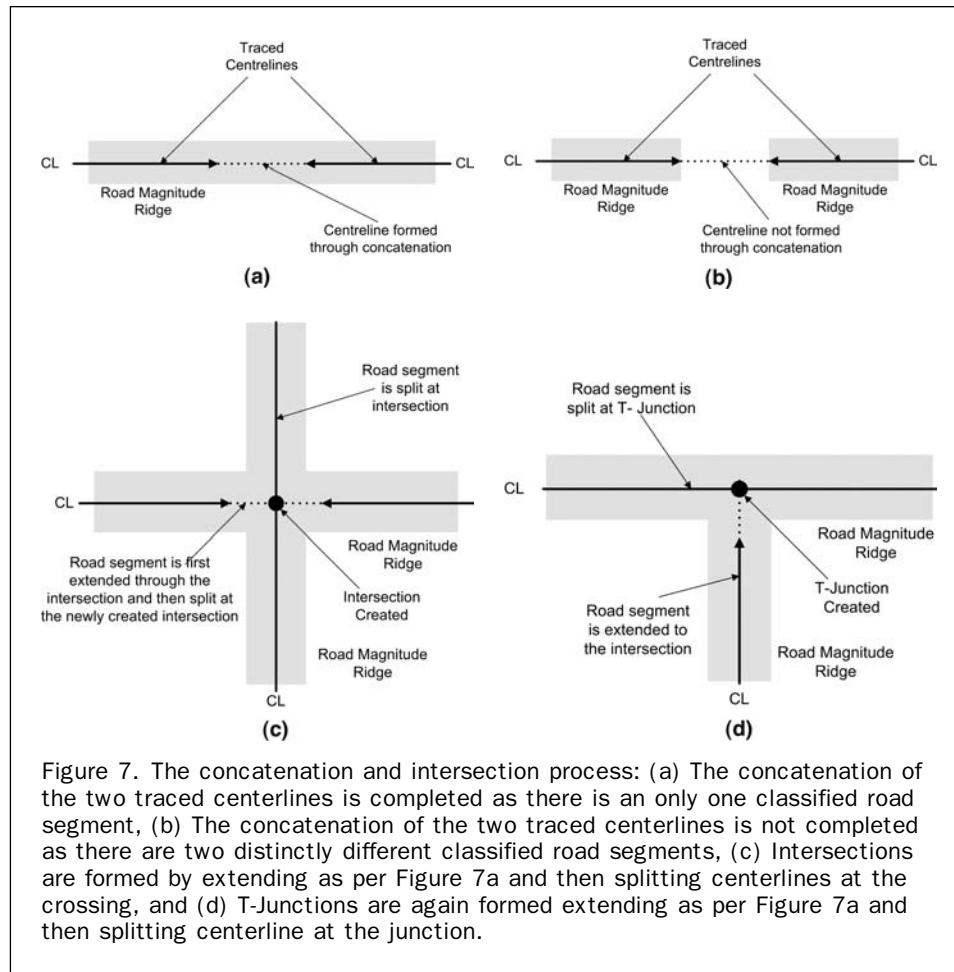


Figure 7. The concatenation and intersection process: (a) The concatenation of the two traced centerlines is completed as there is an only one classified road segment, (b) The concatenation of the two traced centerlines is not completed as there are two distinctly different classified road segments, (c) Intersections are formed by extending as per Figure 7a and then splitting centerlines at the crossing, and (d) T-Junctions are again formed extending as per Figure 7a and then splitting centerline at the junction.

park and private roads (driveways and roads leading to car parks) were not. This left the obvious problem of car parks and private roads which were not considered as part of the ground truth. Unfortunately, it is very difficult to distinguish automatically a car park from a road as they have similar properties. Although an attempt to remove car parks was made, the quality numbers presented can only be considered with this in mind.

### Classification

#### Evaluation of Classification Results

Figure 8 shows the results of the classification using both the Fairfield data set (Figure 8a, 8b, and 8c) and the Yeronga data set (Figure 8d, 8e, and 8f). The manually digitized ground truth images of each area are shown in Figure 8a and 8d as binary images (as negatives). The classification results from the workflow described in the previous section are displayed as binary images of road pixels in the center (as a negative). A visual comparison of the ground truth and binary classified data implies that the classification was successful.

The accuracy of any road extraction technique can be summarized by contemplating the *completeness*, *correctness*, and the *quality* of the detected road network as defined in Heipke *et al.* (1997). In Equations 22, 23, and 24, *TP* denotes the number of true positives, which is the number of pixels found in both the reference and detected data sets. *FN* is the number of false negatives, which is the number of pixels in the reference data set that were not found in the detected data set. *FP* is the number of false positives, which

is the number of pixels that were detected but did not exist in the reference data set. For comparison purposes, the road classification is performed on a pixel-by-pixel basis. *Completeness*, sometimes called recall, is the ratio of the correctly extracted records to the total number of relevant pixels within the ground truth data. *Correctness*, sometimes called precision, is the ratio of the number of relevant pixels extracted to the total number of relevant and irrelevant pixels retrieved. *Quality* is the ratio of the number of relevant pixels extracted to the total number of relevant pixels and detected irrelevant pixels retrieved.

$$Completeness = \frac{TP}{TP + FN} \quad (22)$$

$$Correctness = \frac{TP}{TP + FP} \quad (23)$$

$$Quality = \frac{TP}{TP + FP + FN} \quad (24)$$

The spatial distribution of the *TP*, *FP*, and *FN* pixels along with the True Negative *TN* pixels are displayed in Figure 8c and 8f in light grey, black, dark grey, and white, respectively. A perusal of the spatial distribution of Fairfield (Figure 8c) reveals that the majority of the *FP* detections correspond to car parks while the majority of *FN* detections have occurred at the ends of detected roads or on the edge of the image. These road components are disconnected from the road network and exist in small sections due to overhanging

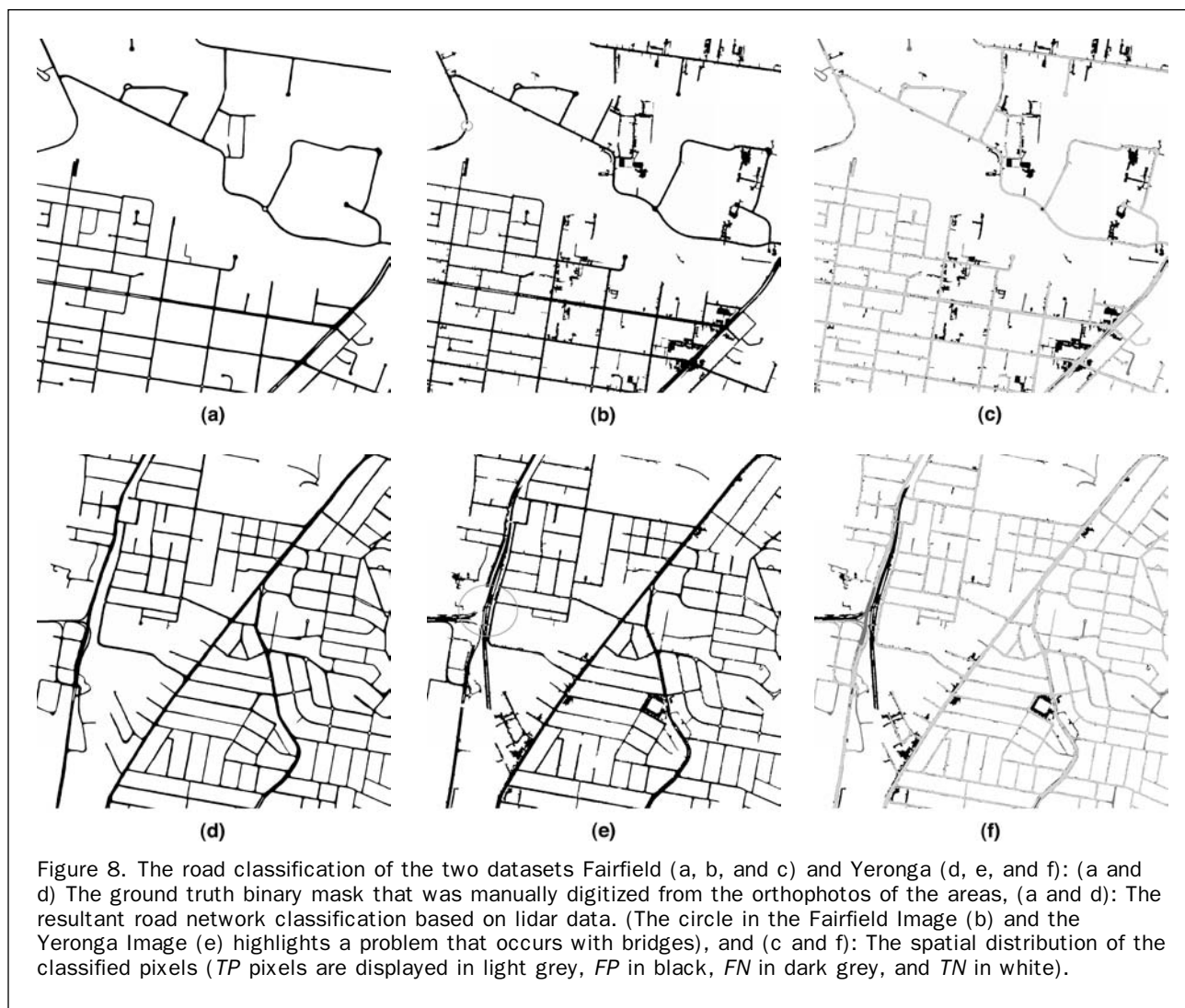


Figure 8. The road classification of the two datasets Fairfield (a, b, and c) and Yeronga (d, e, and f): (a and d) The ground truth binary mask that was manually digitized from the orthophotos of the areas, (a and d): The resultant road network classification based on lidar data. (The circle in the Fairfield Image (b) and the Yeronga Image (e) highlights a problem that occurs with bridges), and (c and f): The spatial distribution of the classified pixels (*TP* pixels are displayed in light grey, *FP* in black, *FN* in dark grey, and *TN* in white).

trees and were removed when small, disconnected components were removed. The spatial distribution of Yeronga (Figure 8f) reveals that there are still some *FP* detections that correspond to car parks, however, the major occurrence is actually a railway that appears on the western side of the image. The railway is naturally not included in the ground truth data, but the rail network in this area has similar properties to the road network, and as a consequence is detected. The majority of *FP* detections that occurred have again been caused by similar factors to those in the Fairfield data set with one major exception. A large bridge was misclassified as “not road” in the western portion of the image.

Table 1 shows the *completeness*, *correctness*, and *quality* numbers obtained for our road classification method when

TABLE 1. THE ACCURACY OF THE ROAD NETWORK CLASSIFICATION IS SUMMARISED IN TERMS OF CORRECTNESS, COMPLETENESS, AND QUALITY

	Completeness	Correctness	Quality
Fairfield	0.88	0.67	0.61
Yeronga	0.79	0.80	0.66

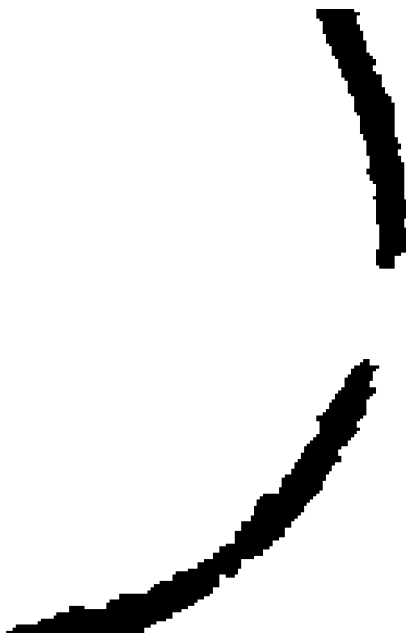
applied to the Fairfield and Yeronga test data. The numbers for *TP*, *FP*, and *FN* were obtained by counting the number of correctly/incorrectly classified pixels in the binary road image as compared to the binary reference image.

Unfortunately, the workflow classification method cannot be directly compared to other algorithms on the same area as data and information was not available. The results from Wiedemann and Hinz (1999) highlight the necessity only to compare algorithms on the same test site, as the same algorithm was run on two different test sites producing substantially different quality results.

There were several problems encountered during the classification phase of the algorithm. One of the problems was the detection of elevated roads and bridges. Our road network model assumed that roads lie on the DTM. As bridges by nature do not lie on the DTM, they are correctly classified in terms of our model assumptions but ultimately incorrect with respect to the reference data. Figure 9 displays an enlarged portion from the northwest corner of the surveyed area displayed by a circle in Figure 8b. The bridge can be seen from the enlarged orthophoto and a row of trees can be seen delineating a creek that runs approximately East to West through the image. The road to be detected in the image runs approximately North to South. At the



(a)



(b)

Figure 9. The Bridge Problem (The area displayed by the circle in Figure 8b): (a) The bridge area is displayed in an enlarged orthophoto for clarity, and (b) The bridge is missing from the classification due to the fact that the bridge is above the DTM.

intersection of the road and the creek, there is a bridge which has clearly not been detected in the enlarged binary classification image seen in Figure 8b and Figure 9b.

Another problem is car parks. Due to the industrial nature of sections of the test data, in particular the Fairfield data set, an attempt to remove some car parks was made. Unfortunately, it is very difficult to distinguish a car park

from a road. Some car parks (not all) are very wide, and thus by defining a maximum acceptable road width prior to processing, it is possible to remove larger car parks from the current binary image. However, the majority of car parks still remain. The quality results will reflect this difference and improved results are expected in non-industrial areas.

#### Algorithm Sensitivity to DTM Noise

In order to determine the effects of DTM noise on the classification algorithm,  $\Delta h_{max}$  from Equation 3 was varied between 0 and 5 meters and the classification of the roads from each data set was repeated. The results of these trials are shown graphically in Figure 10 where the *correctness*, *completeness*, and *quality* are plotted against the threshold. The obtained *correctness* values for both datasets were very close to 1.0 when the threshold was initially set close to 0 m. The *correctness* values decreased as the threshold increased. The *completeness* values showed the opposite trend, i.e., the value increased as the threshold increased. The peak of the *quality* value was expected to show the optimal threshold for each data set. The *quality* value peaks for Fairfield and Yeronga were observed at 0.03 m and 0.35 m, respectively. Although these peak values are significantly different, a closer look at both graphs reveals that only slight variations in the *quality* values are observed between threshold values of approximately 0.05 m and 0.5 m. This suggests that the algorithm is not extremely sensitive to DTM noise, provided the threshold  $\Delta h_{max}$  is chosen appropriately, i.e., between 0.05 m and 0.5 m. It also reinforces the selection of  $\Delta h_{max}$  which was chosen to be 0.3 m for both datasets.

#### Vectorization

As previously described, road information such as the centerline can be obtained from a pre-classified binary image. The result of the convolution of the pre-classified binary images for both Fairfield (Figure 8b) and Yeronga (Figure 8e) with the PCD (Figure 4) is shown in Figure 11. Figure 11a and 11c display the resultant magnitude image for Fairfield (Figure 11a) and Yeronga (Figure 11c) with the highest values being displayed as white and the lowest as black. The ridges within the image correspond to the estimated road centerlines. Figure 11b and 11d display the resultant phase image for Fairfield (Figure 11b) and Yeronga (Figure 11d) and clearly show a difference in the phase at the centerline of the road and the road edge.

The centerlines of the road image are represented by white ridges running through the magnitude images. The corresponding direction of the line can be seen in the phase image. Enlargements of the convolution results are displayed in Figure 12 to highlight the centerline and orientation estimation.

The centerline of the binary classified image has been detected using the convolution and can be clearly seen in Figure 12a. The image shows the centerline of a curved suburban road being extracted from the Fairfield data set. The other image shows an enlarged magnitude image centred over an intersection overlaid with arrows indicating the direction of half the phase. It can be clearly seen that at the centerline the half phase value is along the line, while at the road edges the half phase is  $\pm 90^\circ$  out from the value at the line. The location of both images displayed in Figure 12 is shown by the squares in Figure 11a.

Vectorization of the convolution results was performed by the tracing algorithm as previously described. The results of the vectorization process can be seen in Figure 13. The centerline vectors of both data sets appear visually to be a good approximation of the road network. In areas where there are very few car parks the vectors are smooth,

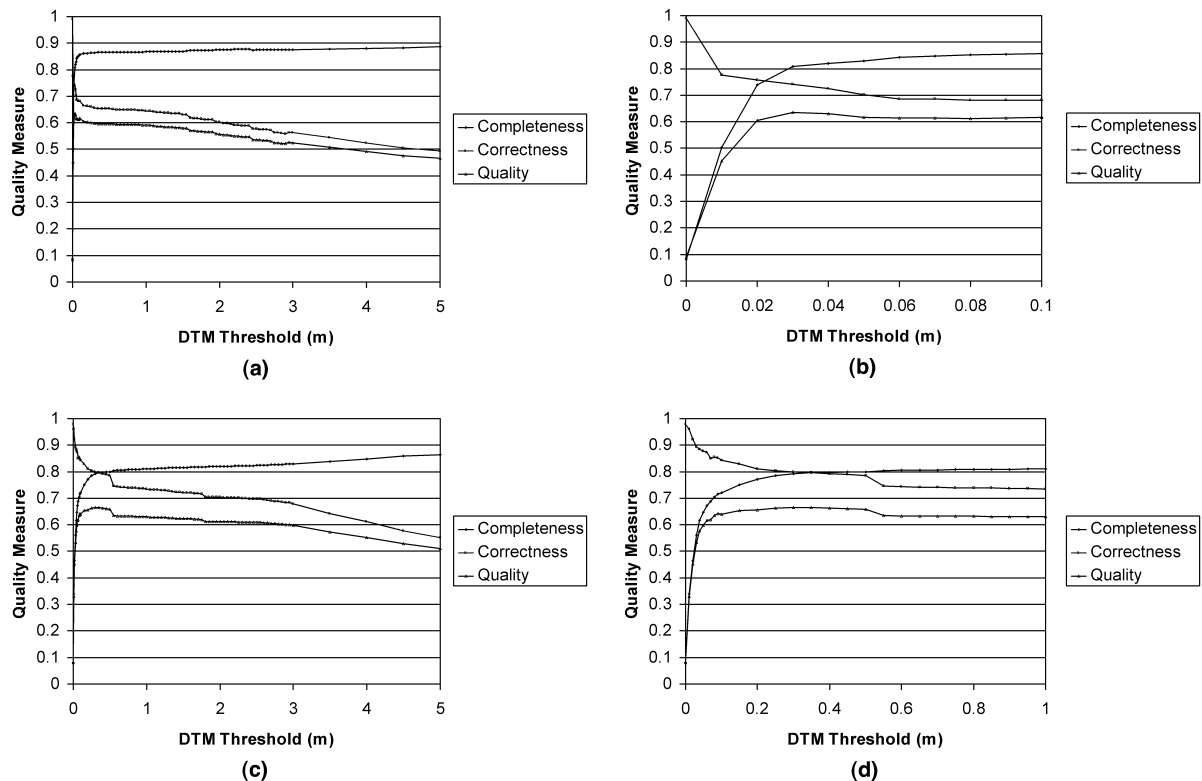


Figure 10. Completeness, correctness and quality measures for Fairfield (a and b) and Yeronga (c and d) for the classification of roads with different DTM thresholds ( $\Delta h_{max}$ ) applied to the DTM to demonstrate the algorithm sensitivity to DTM accuracy. (a and c) Thresholds were varied from 0 to 5 metres for both data sets. (b and d) Enlargements of the areas of interest to allow better interpretation of the results.

predictable, and continuous. This is very similar to our original road network model assumptions. The results from the Fairfield dataset, Figure 13a and 13b, show that roads that have successfully been classified as dual carriageways have also been vectorized as dual carriageways in these areas. The curvilinear road that runs diagonally from the northwest corner to the eastern edge of the image is quite smooth and characterizes the classified image well. The grid pattern in the southwestern corner is very good. There are several regions within Fairfield that appear to be quite noisy, in particular, the area in the southeast of the image. The behavior of the tracing algorithm is predominately as expected, although some anomalies could be seen. In the southwest corner of Figure 13a there are two intersections that are represented as a set of two 90° bends. In these instances, the ridge in the magnitude image was more dominant towards the entering road rather than the continuing road. Another problem also seen in the southwest corner of the image is the tracing of roundabouts. Due to the nature of the zeroing in the tracing algorithm, only three quarters of the roundabout has been effectively traced. This apparent noise is due to the presence of many car parks in the area. The vectorization algorithm has attempted to fit a number of road centerlines to the car parks. A close inspection reveals that the vectors resemble the car park areas, however depiction of a centerline in these cases is almost an arbitrary task. It can be seen in the magnitude image displayed in Figure 11 that the bridge in the northwest corner of the image (indicated by the white circle) has still yielded a

continuous ridge, albeit reduced in magnitude. The ridge is still present as the gap in the initial classification was less than the size of the PCD used. Thus, the tracing algorithm correctly vectorized this area, and information that was lost in the initial classification has been recovered.

The results from the Yeronga dataset, Figure 13c and 13d, again visually show a good approximation of the road network. The grid pattern in general is characterized by the vectorization process. There are several roundabouts in the image that have again caused problems. The large bridge on the western edge of the image (shown by a circle in Figure 8e and Figure 11c) has not been recovered as the bridge length was longer the proximity measure used to concatenate road segments. This is the opposite result to that observed for the bridge highlighted in the Fairfield data set which was recovered.

The edges of the roads are also displayed in Figure 13b and 13d for both datasets. In areas of relatively low noise, the road edges have been calculated consistently. The image appears similar to what would be expected from the output of a standard edge detection algorithm, such as Canny, with one difference. The magnitude image was created from the convolution of the binary classified image and the PCD, thus the width of the road at any particular point along a road is effectively being smoothed over the local area of the PCD itself. Smoother edges are obtained along areas of clearly defined roads. In areas where roads are less clearly defined, e.g., with car parks or road junctions, the resultant magnitude is less representative of the road due to a saturation effect within the PCD. The width in

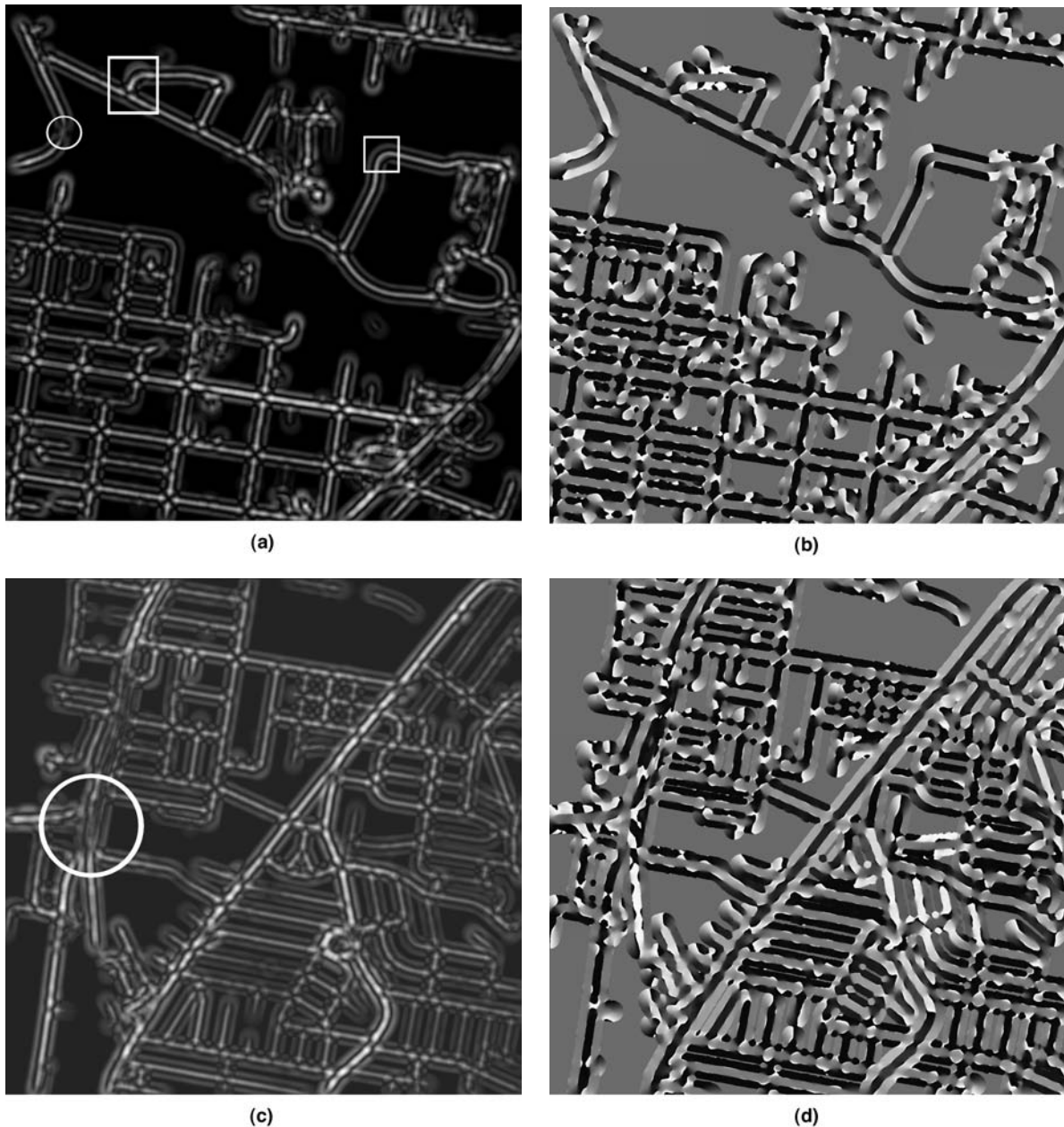


Figure 11. Results of the convolution with the PCD for Fairfield (a and b) and Yeronga (c and d): (a and c) The resultant magnitude images derived from the convolution with the highest values being displayed as white and the lowest as black. (The circle in the Fairfield magnitude image indicates the region described by Figure 9 which demonstrates a bridge being missed by the classification. The two squares in the Fairfield magnitude image indicate the areas discussed in Figure 12), and (b and d) The resultant phase images from the convolutions.

these areas appears to be less reliable, so a smoothed road width is used along a road segment, thus ensuring a consistent width along the segment.

In order to quantify the vectorization results, the evaluation methods described in Wiedemann (2003) have been adopted. One major difference to the methods of Wiedemann (2003) was that the reference or ground truth data was available only in raster format and not vector format. The vectorized road image is hence matched to an estimated ground truth centerline position based on a predetermined

buffer width. The estimated reference centerline is obtained by detecting the two edges of the binary image along the line formed by the width at each vertex. The edges are used to calculate a matching centerline position and width for each extracted vertex. The quality measures for all centerline points, road segments, and road widths are then calculated.

In order to quantify the vectorization quality, the buffer width for the simple matching technique defined by Wiedemann (2003) is initially set. *Completeness* describes how well reference data is explained by the extracted data



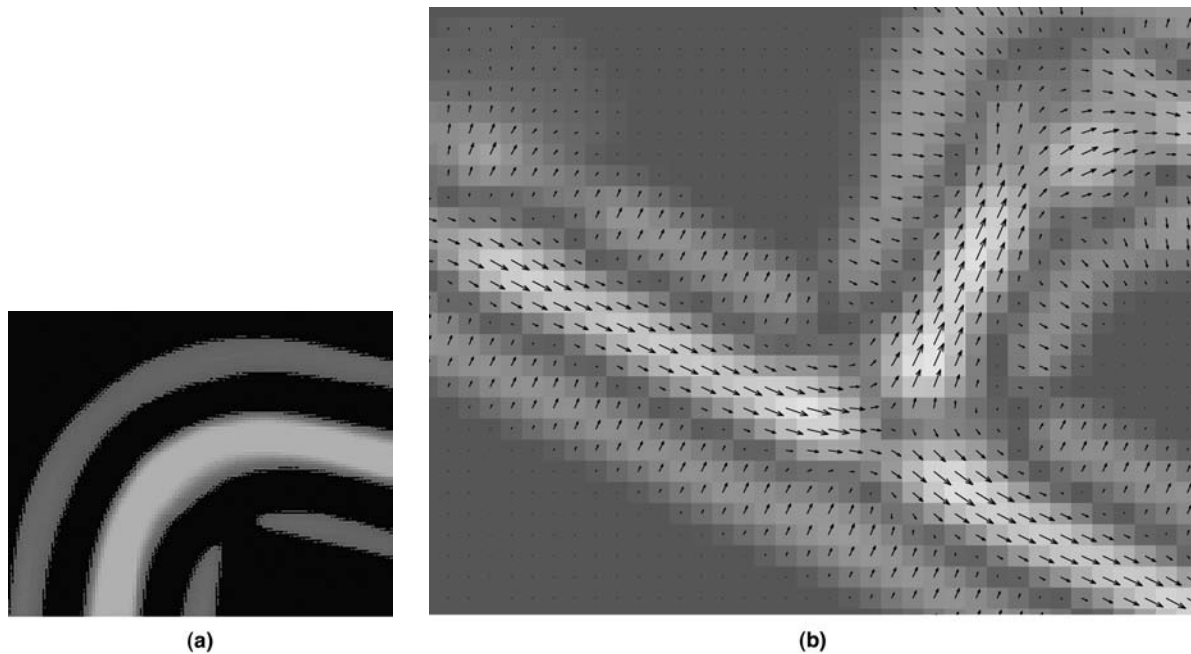


Figure 12. Enlarged results of the convolution with the PCD indicated by squares in Figure 11: (a) The enlarged area displays a windy suburban road that has had the road centerlines detected using convolution with the PCD. (For clarity, the highest values are being displayed as white and the lowest as black), and (b) The magnitude image is overlaid with arrows indicating the direction of half the phase at an enlarged intersection. (Note: The vector length is proportional to the magnitude.)

and is determined by the ratio of the *length of matched reference data* to the *total length of reference data*. *Correctness* represents the percentage of correctly extracted road data and is determined by the ratio of the *length of matched extraction* to *total length of matched extraction*. The centerline root mean square (RMS) of all points expresses the quadratic mean distance between the matched extracted and matched reference points. The centerline RMS of all road segments considers the RMS values for each individual road segment and expresses the quadratic mean distance of these results. The width RMS of all points and road segments are calculated in a similar manner to the centerline quality values but consider the difference in extracted and reference widths as opposed to the centerlines. In order to remove the influence of outliers, the buffer width was also used to match the widths. Detected intersections were manually classified against the reference data into *TP*, *FP*, and *FN* detections. Topological *completeness* and *correctness* are then calculated using Equations 22 and 23. The accuracy of the road vectorization process is summarized in Table 2. The final evaluation of the vectorized road network is promising, and the results resemble the classification results. This suggests that the vectorization process was performed well, and that improvements need to be made to the classification algorithm in order to achieve better final quality values. The topological evaluation of the network extraction suggests that improvement in the current network model can be made.

Figure 14 shows enlargements of the vectorized results overlaid on an aerial image of the area in order to demonstrate the overall geometric quality of the classification and vectorization. The enlargements have been taken from the Fairfield data set and are depicted in Figure 13 by a rectangle.

## Conclusions

This paper describes an effective method for the classification of roads from lidar data using both the intensity and range data. The results achieved appear encouraging although a direct comparison could not be made to other published methods that use other data sources such as aerial photography. The classification is performed on two large urban areas. Both urban and rural traits are present in the data and the algorithm presented appears to work well in all areas. However, car parks pose some problems. The presence of many car parks and private roads has reduced the achieved *correctness* value due to the high presence of false positives as car parks and private roads are not considered in the ground truth data. *Completeness* values of 0.88 and 0.79 and *correctness* values of 0.67 and 0.80 were achieved for the classification phase of the process.

Centerline extraction is performed by the convolution of the binary road image with a complex kernel which enables extra information to be extracted from the classified image. Both the width and direction of the road at the centerline are obtained from the convolution. The direction of the road helps facilitate the successful vectorization of the magnitude image by way of the tracing algorithm described. The vectorization of any classified road network captured in a binary image can be achieved using the PCD method. The vectorization of the classified results yielded RMS values of 1.56 m and 1.66 m, *completeness* values of 0.84 and 0.81, and *correctness* values of 0.75 and 0.80 for two different data sets.

Future work is to be focused on the classification of the lidar data. The removal of car parks from the classified road image would greatly improve the classification results. The application of several PCD's with different radii within a



Figure 13. The final vectorized road images for Fairfield (a and b) and Yeronga (c and d): (a and c) The vectorized road centerlines, and (b and d): The vectorized road edges as calculated from the road centerline and the detected road width.

TABLE 2. THE ACCURACY OF THE ROAD VECTORIZATION PROCESS IS SUMMARISED IN TERMS OF COMPLETENESS, CORRECTNESS, AND THE RMS VALUES AS DESCRIBED IN WIEDEMANN (2003)

	Fairfield	Yeronga
Completeness	0.84	0.81
Correctness	0.75	0.80
RMS of all points – Centerline (m)	1.70	1.77
RMS of all road segments – Centerline (m)	1.56	1.66
RMS of all points – Width (m)	1.66	1.55
RMS of all road segments – Width (m)	1.48	0.93
Topological Completeness	0.87	0.81
Topological Correctness	0.73	0.89

hierarchical framework during the vectorization stage may improve the vectorization of the road edges and allow a more robust determination of the road width. An improved road model at the intersection of roads may also improve the final road network.

### Acknowledgments

This work was supported by the Australian Research Council (ARC) under Linkage Project LP0230563 and Discovery Project DP0344678. The Fairfield and Yeronga data sets were provided by AAMHatch, Fortitude Valley, QLD 4006, Australia ([www.aamhatch.com.au](http://www.aamhatch.com.au)).



(a)



(b)

Figure 14. The road centerline and edges are overlaid on an aerial image to demonstrate the overall geometric quality of the vectorized results (The area displayed by the rectangle in Figure 13a). (a) The overlaid vectorized road centerlines, and (b) The overlaid vectorized road edges.

## References

- Akel, N.A., O. Zilberstein, and Y. Doytsher, 2003. Automatic DTM extraction from dense raw LIDAR data in urban areas *Proceedings of the FIG Working Week*, URL: <http://www.fig.net/pub/fig.2003/program.htm> (last date accessed: 04 January 2007).
- Alharthy, A., and J. Bethel, 2003. Automated road extraction from LIDAR data, *Proceedings of the ASPRS Annual Conference*, Anchorage, Alaska, unpaginated CD-ROM.
- Atherton, T., and D. Kerbyson, 1999. Size invariant circle estimation, *Image and Vision Computing*, 17:795–803.
- Auclair-Fortier, M.-F., D. Ziou, C. Armenakis, and S. Wang, 2001. Survey of work on road extraction in aerial and satellite images, *Canadian Journal of Remote Sensing*, 27(1):76–89.

- Clode, S., P. Kootsookos, and F. Rottensteiner, 2004a. The automatic extraction of roads from LIDAR data, *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Science*, Vol. XXXV-B3, pp. 231–236.
- Clode, S.P., E.E. Zelniker, P.J. Kootsookos, and I.V.L. Clarkson, 2004b. A phase coded disk approach to thick curvilinear line detection, *Proceedings of EUSIPCO*, Vienna, Austria, pp. 1147–1150.
- Duda R., and P.E. Hart, 1972. Use of the Hough Transform to detect lines and curves in pictures, *Communications of the Association of Computing Machines*, 15(1):11–15.
- Hatger, C., and C. Brenner, 2003. Extraction of road geometry parameters from laser scanning and existing databases, *Proceedings of the Workshop: 3-D Reconstruction from Airborne Laserscanner and InSAR Data*, *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Science*, Vol. XXXIV, Part 3/W13, Dresden, Germany.
- Heipke, C., H. Mayer, C. Wiedemann, and O. Jamet, 1997. Evaluation of automatic road extraction, *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXII-3/2W3, pp. 47–56.
- Hinz, S., 2004. Automatic road extraction in urban scenes, *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Science*, Vol. XXXV-B3, pp. 349–354.
- Hinz, S., and A. Baumgartner, 2003. Automatic extraction of urban road networks from multi-view aerial imagery, *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(1–2):83–98.
- Hinz, S., and C. Wiedemann, 2004. Increasing efficiency of road extraction by self-diagnosis, *Photogrammetric Engineering & Remote Sensing*, 70(12):1433–1440.
- Hu, X., and C.V. Tao, 2003. Automatic extraction of main-road centerlines from high resolution satellite imagery based on perceptual grouping, *Proceedings of the ASPRS Annual Conference*, Anchorage, Alaska, unpaginated, CD-ROM.
- Hu, X., C.V. Tao, and Y. Hu, 2004. Automatic road extraction from dense urban area by integrated processing of high resolution imagery and LIDAR data, *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Science*, Vol. XXXV-B3, pp. 288–292.
- Huber, R., and K. Lang, 2001. Road extraction from high-resolution airborne SAR using operator fusion, *Proceedings of the International Geoscience and Remote Sensing Symposium*, Vol. 5, pp. 2813–2815.
- Kraus, K., 2002. Principles of airborne laser scanning, *Journal of the Swedish Society of Photogrammetry and Remote Sensing*, 1:53–56.
- Lee, H.Y., W. Park, H.-K. Lee, and T.-G. Kim, 2000. Towards knowledge-based extraction of roads from 1 m resolution satellite images: *Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation*, Austin, Texas, pp. 171–176.
- Mena, J.B., 2003. State of the art on automatic road extraction for GIS update: A novel classification, *Pattern Recognition Letters*, 24(16):3037–3058.
- Priestnall, G., M.J. Hatcher, M.D. Morton, S.J. Wallace, and R.G. Ley, 2004. A framework for automated extraction and classification of linear networks, *Photogrammetric Engineering & Remote Sensing*, 70(12):1373–1382.
- Rieger, W., M. Kerschner, T. Reiter, and F. Rottensteiner, 1999. Roads and buildings from laser scanner data within a forest enterprise, *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXII/3-W14, La Jolla, California, pp. 185–191.
- Rottensteiner, F., J. Trinder, S. Clode, and K. Kubik, 2003. Building detection using LIDAR data and multispectral images, *Proceedings of DICTA*, Sydney, Australia, Vol. 2, pp. 673–682.
- Steger, C., 1996. Extracting curvilinear structures: A differential geometric approach, *Proceedings of the Fourth European Conference on Computer Vision*, Vol. 1064, pp. 630–641.
- Vosselman, G., 2002. On the estimation of planimetric offsets in laser altimetry data, *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Science*, Vol. XXXIV/3A, pp. 375–380.

- Weidner, U., and W. Förstner, 1995. Towards automatic building reconstruction from high resolution digital elevation models, *ISPRS Journal of Photogrammetry and Remote Sensing*, 50(4): 38–49.
- Wiedemann, C., 2003. External evaluation of road networks, *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXIV-3/W8, pp. 93–98.
- Wiedemann, C., and S. Hinz, 1999. Automatic extraction and evaluation of road networks from satellite imagery, *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXII, pp. 95–100.
- Zhang, C., 2003. *Updating of Cartographic Road Databases by Image Analysis*, Ph.D. dissertation, Mitteilungen Nr 79 of the Institute of Geodesy and Photogrammetry at ETH, Zurich, Switzerland.
- Zhu, P., Z. Lu, X. Chen, K. Honda, and A. Eiumnoh, 2004. Extraction of city roads through shadow path reconstruction using laser data, *Photogrammetric Engineering & Remote Sensing*, 70(12):1433–1440.