Available online at www.sciencedirect.com
SCIENCE DIRECT®
Information Fusion 6 (2005) 257–273
INFORMATION FUSION
www.elsevier.com/locate/inffus

An integrated system for automatic road mapping from high-resolution multi-spectral satellite imagery by information fusion

Xiaoying Jin, Curt H. Davis *

Department of Electrical and Computer Engineering, University of Missouri – Columbia, Columbia, MO 65211, USA

Received 29 September 2003; received in revised form 15 June 2004; accepted 17 June 2004
Available online 19 July 2004

Abstract

Here we present an integrated system for automatic mapping of urban and suburban roads from high-resolution satellite imagery. Road extraction strategies can be enriched and improved by multi-source data fusion. Fusion can occur by combining road extraction results from different image sources, or by applying multiple detectors to a single source image and then fusing the multi-detector road extraction results. Our proposed system models roads differently for urban vs. suburban environments. In suburban areas, roads are modeled as curvilinear and homogeneous regions with nearly constant width. Roads are extracted by integrating the output of two road detectors followed by a road tracker. The first detector is based on fine-scale image segmentation and grouping where road centerlines are identified and extracted using shape and structure information from the segments. This detector generates a very reliable road extraction but is incomplete as many road segments are not captured. A multi-scale curvilinear structure detector based on differential geometry is used as a complementary detector. This detector utilizes a fuzzy decision algorithm based on a 1D road profile model. The final suburban road centerlines are generated by integrating the results from the two detectors using an optimum path search algorithm. In many dense urban areas in the US, road networks are composed primarily of straight lines that form a grid structure. Here we use directional morphological filtering to mask out dark and bright structures shorter than a city block. We then use a spatial signature weighted Hough transform to generate a road grid hypothesis. The modified Hough transform incorporates spatial information derived from surrounding image pixels and can be applied to multi-spectral images. Each piece of the hypothesized road grid is then verified using homogeneity, shape, and vegetation information from the local surrounding area. Missing pieces of the road network are added using a road tracker based on profile matching. The integrated road extraction system based on the fusion of multi-detector results is tested using IKONOS multi-spectral imagery of the City of Columbia, Missouri. Evaluation of the extracted road networks using representative test sites show completeness values that range between 70% and 86% and correctness values that range between 70% and 92%.

Keywords: Automatic road extraction; High-resolution satellite imagery; Perceptual grouping; Hypothesis and verification; Multi-detector fusion

1. Introduction

The recent availability of commercial high-resolution satellite imaging sensors such as IKONOS and QuickBird provide an important source for urban remote sensing applications. The high spatial resolution of the imagery reveals very fine details in urban areas and greatly facilitates the recognition and extraction of urban-related features [7,8,11,23,24]. Launched in September 1999, IKONOS was the first commercial high-resolution satellite. IKONOS collects 1 m panchromatic (PAN) and 4 m multi-spectral (MS) imagery. With its high spatial resolution and geometric accuracy, it is possible to identify fine-scale features such as individual roads and buildings in the urban environment and also provide very accurate geodetic coordinates. Accurate and up-to-date road network information is essential for

* Corresponding author. Tel.: +1-573-884-3789; fax: +1-573-882-0397.
E-mail addresses: xje4e@mizzou.edu (X. Jin), davisch@missouri.edu (C.H. Davis).

1566-2535/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.inffus.2004.06.003

geographic information system (GIS) databases, transportation and urban planning, automated road navigation, and emergency response applications. Many GIS road network data layers are spatially inaccurate and/or out of date. In addition, rapidly changing urban environments precipitate the need for frequent updates or revisions. Since manual extraction of road networks from imagery is very time consuming, automated methods have the potential to improve the speed and utility for road mapping and are therefore highly desirable.

Research on road extraction from aerial and satellite images can be traced back to the seventies [22]. Early approaches for road detection were developed using both low and high-resolution aerial photographs [4,22]. In low and medium resolution images, roads were modeled as continuous and smooth lines in the imagery. Here, the road extraction was often equated with linear feature extraction [4,27,30]. In high-resolution images, geometric quantities such as structure and shape play a crucial role in road recognition, where roads are often modeled as continuous and elongated homogeneous regions with nearly constant width [7,11]. Road extraction strategies are usually classified into two categories according to the degree of human interaction: semi-automated and automated extraction. For semi-automated road extraction, initial seed points, sometimes with directions, must be provided to an algorithm that attempts to connect the seed points using various search path criteria. For automated extraction, salient roads, called road seeds, must be detected automatically and tracked or linked to complete the road network.

Although a significant amount of research exists on road extraction, most of this research has focused on rural and/or forested areas. Much less research has been performed for complex urban–suburban areas. Moreover, quantitative performance metrics for road extraction accuracy are seldom reported for urban–suburban areas. Automated road extraction is a complex and difficult task and, as such, any single-source approach is likely to be inadequate for routine application. Robust and reliable road extraction is better achieved by fusion of multi-source information. Fusion can occur by combining road extraction results from different image sources (SAR, multi-spectral, hyperspectral, satellite, airborne, etc.) or by applying multiple detectors to a single source image and then fusing the multi-detector road extraction results. In addition, significantly different approaches should be exploited depending on the contextual setting of the image. For example, in dense urban areas, road networks in US cities are often characterized by a series of parallel and orthogonal straight lines grouped to form a grid structure, whereas in suburban areas, roads are mostly curvilinear structures having a constant width and have a much less rigid macroscopic spatial structure. Accordingly, different approaches are normally required for each contextual setting. In addition, multiple detectors or methods are required for robust extraction because of the diversity of image illumination, road material, and scene context. If a portion of the road network does not fit a given model, it cannot be extracted with high accuracy. In order to cope with the high complexity of real scenes, integrating the power of multiple road extraction approaches is needed to improve the reliability and robustness of the extraction results [5].

Although much research has been done for road extraction from a diversity of image sources [18], only a few papers have investigated the fusion of different information sources for road extraction. Early information fusion work was performed on low and high-resolution aerial imagery. Fischler et al. [4] partitioned image operators into two classes: Type I operators have low false alarm rates but may often miss correct targets; whereas Type II operators provide precise local feature characterization but may falsely identify non-targets. Results from both Type I and II operators can be combined by either a graph search or dynamic programming. Fischler et al. [4] used linear structure detectors applied to low-resolution images for both Type I and II operator classes. McKeown et al. [16] used two low-level cooperative methods for road extraction from high-resolution aerial imagery: road surface texture correlation and road-edge following. Intermediate-level processes were used to monitor the state of the low-level feature extraction methods and to perform method evaluations. If one tracking method failed, it could be restarted from the model generated by the other successful tracker.

Information fusion strategies for road extraction have recently been explored using SAR images. Tupin et al. [30] used two road detectors. One detector was based on a ratio-edge detector and the other detector used the correlation between two populations of pixels. The results from the two detectors were fused by using a mathematical operation called associative symmetrical summation. In the second global step, the candidate road segments were identified by introducing contextual knowledge by defining a Markov random field (MRF) on a set of segments. Most recently, Dell'Acqua and Gamba [2] used multiple detectors in SAR data for road extraction. They used hard or fuzzy "AND" and "OR" operators compared with majority voting and log-opinion pool to fuse three different operators for urban street mapping: fuzzy connectivity weighted Hough transform (FCWHT), fuzzy pyramidal rotation transform (FPRT), and fuzzy shortest path extraction (FSPE). Hinz and Baumgartner [9] performed automatic extraction of urban road networks from multi-view aerial imagery. For global context, they identified urban, forest, and rural regions to allow specific road models and extraction strategies to adapt to the contextual situation. Lane

segments and vehicles extracted from multiple images were projected onto a Digital Surface Model (DSM) and merged based on the objects' self-diagnosed confidence values to fuse the results from the multi-view aerial images. Very little work has been done to date on information fusion for road extraction from high-resolution optical satellite imagery. Haverkamp et al. [8] used an automated method to extract straight-line, gridded road networks that was then followed by a semi-automated road tracker tool to approximate curvilinear roads and fill in missing pieces of the road network. Their automated extraction results achieved 49–66% road extraction completeness for various test sites [8].

In this paper we present an integrated system for automated urban and suburban road mapping from high-resolution multi-spectral satellite images. The system runs automatically without the need for preselected points, although some algorithm parameters must be set. Roads are modeled differently depending on the contextual environment. First, dense urban areas are distinguished from suburban areas using vegetation density. In suburban areas, roads are modeled as curvilinear and homogeneous regions with nearly constant width. Here, roads are extracted by integration of two complementary road detectors. A road tracker is then used to follow and connect missing parts of the road network. In dense urban areas, the road extraction is based on street grid hypothesis and verification [21]. The integrated road extraction system based on the fusion of multi-detector outputs is tested using IKONOS imagery of the City of Columbia, Missouri. In this study, the PAN data are fused with the MS data using a color normalization method [33] implemented in ENVI 3.5 to generate a four-band pan-sharpened multi-spectral (PS-MS) image with 1 m spatial resolution. Quantitative evaluations of the extracted road network from representative test sites are reported. By fusing the results from multiple detectors, the quality and accuracy of the road network extraction is greatly improved. Extracted road network completeness values range between 70% and 86% and the correctness values range between 70% and 92%.

The remainder of this paper is organized as follows. In Section 2, we first present an overview of the integrated system. Next, the preprocessing techniques and the method used to differentiate dense urban areas from suburban areas are described. In Section 3, the road extraction methods used in suburban areas and the corresponding fusion strategy are described. In Section 4, the dense urban area methodology is described. Here, the road extraction utilizes grid hypothesis and verification and the results are then fused with the extraction results from the surrounding suburban areas. In Section 5, quantitative results and comparisons are presented. Finally, conclusions are given in Section 6.

2. Integrated system overview

Roads typically have widths ranging between 8 and 30 m. Roads may be made of different materials, such as asphalt, concrete, gravel, etc. Tall buildings and/or trees along the roads can cast shadows. Spectral mixing and overlap with other urban land cover classes is also a problem (e.g. bare soil, parking lot, etc.). Consequently, the spectral characteristics of roads have large variance and are greatly affected by their context. This is especially true for urban environments, where objects such as buildings and parking lots may have the same spectral response as roads. Because of the complex nature of road extraction in real-world environments, our integrated system incorporates multiple detectors to obtain reliable road vector data.

2.1. Road extraction workflow

The first step in the system is preprocessing. The preprocessing is designed to improve the quality of the image and identify relevant image pixels for further processing. In high-resolution satellite images, a significant amount of detail can be discerned on the roads, such as cars, street markings, etc. These details on one hand can provide additional context information for feature extraction, but they also disturb the spectral characteristics of the overall road feature class we are interested in. Thus, a filtering operation is useful for eliminating or minimizing the effect of these unwanted spectral disturbances. Also, it is beneficial to mask out vegetation and water bodies in the image to reduce computational load and detector confusion. Since roads appear as elongated structures in the image, we use a spatial signature measure to provide shape information for each non-vegetated, non-water pixel to distinguish potential road pixels from other compact impervious surfaces (e.g. parking lots, buildings, etc.). Fuzzy edge information for segmentation grouping (performed later) is also obtained in this step.

As discussed previously, different areas in an urban setting have distinguishing characteristics. Dense urban areas can be distinguished from suburban areas by their low vegetation density and their high density of commercial buildings. In dense urban areas in US cities, roads often appear as straight lines in parallel and orthogonal groups that form a grid structure. In these areas a large number of spectrally similar impervious surfaces exist such as parking lots and commercial buildings. Thus, road extraction is better done using a more global approach. However, suburban areas have a high vegetation density and the roads appear amongst vegetated areas and curvilinear groups of single-family houses. Here the roads are fairly homogeneous and elongated segments with nearly constant width and are relatively easy to distinguish from their surrounding environment based on local spectral information.

In our integrated approach, we adopt different road extraction methodologies, e.g. global vs. local, depending on the contextual setting, e.g. dense urban vs. suburban. In dense urban areas, the Hough transform is a good choice for the localization of straight lines. However, the original Hough transform can only be applied to binary data. Here we propose a spatial signature weighted Hough transform (SSWHT). This modified Hough transform exploits the directional similarity of pixels in a surrounding area and can be applied to multispectral imagery. In dense urban areas, there are many dark or bright structures such as buildings or parking lots. They are usually shorter than the distance of a city block, which we define here to be the average separation distance of road centerlines from two consecutive and parallel main streets in the gridded parts of the city. If many of these features align along a line, there will be a significant response in the polar plane. Thus, we use directional morphological operations to filter out these structures. After generating a road grid hypothesis based on the SSWHT, each piece of the hypothesized road grid is then verified using information from the surrounding local area for each individual road segment.

In suburban areas, we generate salient road centerlines using two complementary detectors. The first detector is based on fine-scale segmentation and grouping. Road seeds are selected from homogeneous segments using shape information. Then perceptual grouping theory is used to group the seeds based on structural information from the seeds. Road centerline selections are then made from the first detector. The detected centerlines are very reliable, but a significant number of road segments are missed. A multi-scale curvilinear structure detector based on differential geometry is used as a complementary detector. The second detector incorporates a fuzzy decision based on 1D road profile modeling. Finally, robust and reliable road centerlines are generated by integrating the results from the two detectors using an optimum path search algorithm.

After extracting road networks from the dense urban vs. suburban environments individually, the results are combined together to generate one vector layer. In the regions between dense urban and suburban areas, some overlap is allowed to let the individual methods work together and make the transition between these two environments seamless. Finally, missing pieces of the road network are added using a road tracker based on profile matching. A flow chart outlining the basic steps of the integrated road extraction system is shown in Fig. 1.

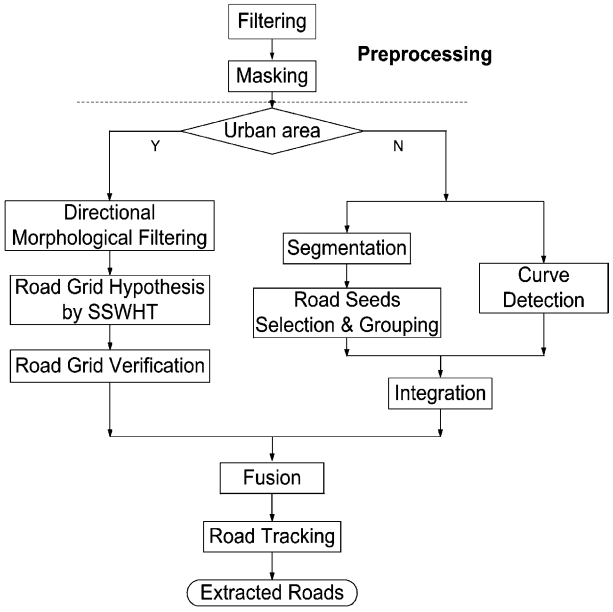

Fig. 1. Processing flow chart for the integrated road mapping system.

2.2. Preprocessing

A series of preprocessing steps are needed to improve image quality and to identify specific image pixels for the further processing. These steps are summarized below.

Filtering. From empirical observation, the typical width of most roads occurs within a range from 8 to 30 m. This corresponds to 8–30 pixels in the 1 m PS-MS IKONOS imagery. A morphological opening operation followed by a closing operation is applied to each channel of the PS-MS data to smooth out cars and street markings widely observed on the roads. The structure element (SE) used was a disc with radius $r = 3$. With a SE at this scale, roads and buildings will not be affected. Small bright structures are smoothed out by the opening operation and small dark structures are smoothed out by the closing operation. Normal opening/closing will distort the shape of structures remaining in the image. Thus, we used morphological opening/closing by reconstruction based on geodesic distance [25,31] to retain the original shape information. After morphological smoothing, a median filter with a 7×7 kernel was then used to further smooth the spectral response within the pixel's local neighborhood.

Masking. To reduce computational load and concentrate future processing on viable regions of interest, vegetated portions of the image are masked out using a threshold applied to the Normalized Difference Vegetation Index (NDVI):

$$\mathrm{NDVI} = \frac{\mathrm{NIR} - R}{\mathrm{NIR} + R}, \tag{1}$$

where NIR is the reflectance value of the near-infrared channel, and R is the reflectance value of the red channel. Vegetation has a high reflectance in the NIR channel and a low reflectance in the R channel. Consequently, vegetated areas with an NDVI higher than 0.3 were masked out.

In addition, we found that some small waterways (creeks) can have similar curvilinear shape characteristics as roads in the PS-MS imagery. We implemented a band math operation to distinguish them from roads. Usually water bodies have a high reflectance in the G channel and a low reflectance in the R channel. Thus, we used

$$\frac{G-R}{G+R} \tag{2}$$

to distinguish water bodies where values higher than 0.15 were masked out.

Spatial signature measure. In shape terminology, a signature is a 1D functional representation of the boundary of a 2D area [6]. We borrowed this concept here to quantify local shape information for each pixel in the multi-spectral image. For each pixel, a search along multiple directions radiating from the current pixel continues until the dissimilarity between the search pixel and the current pixel exceeds a certain level. The Euclidian distance between the spectrums of the pixels was used as the dissimilarity measure. During the tracing, small gaps of a few pixels in length are tolerated. The maximum search length along each direction of the IKONOS imagery was set to be about 200 meters (the suggested value is 1.5∼2 times of the distance of an average city block). To reduce the computational load, directional angles ranging from 0° to 360° were discretized into 72 steps. Please refer to Fig. 2(a) for a graphical representation of the radial image search and Fig. 2(b) for a plot of the spatial signature vs. search direction. In Fig. 2(a), a pixel in the middle of a road segment has two dominant directions that differ by 180°. For each pixel, the maximum sum of the traced lengths along the two opposing directions is used as the length measure and the minimum sum is used as the width measure. The direction with the maximum traced length is recorded as the pixel's primary direction. This approach was adapted from the length-width contextual algorithm in [23] that was used to improve the discrimination between roads and buildings for urban land cover classification of IKONOS PS-MS image data.

Edge detection. The Nevatia–Babu edge operator [20] is used to generate an edge magnitude response rather than the simple Sobel operator. The former operator includes edge detection and thinning. For multi-spectral images, the edge magnitude on each band is obtained individually and then combined using the MAX operation. Thinning is done on the edge magnitude to localize the edge. An iterative thresholding scheme [6] is adopted to determine the edge threshold. Based on fuzzy set theory [12], we use a membership function to represent the degree that a pixel belongs to the edge set. As a result, a fuzzy edge map is produced where each pixel has an edge magnitude in the range of $[0, 1]$. We use a fuzzy edge map rather than a crisp edge map to keep the edge magnitude information for later global use. Please refer to [10] for details.

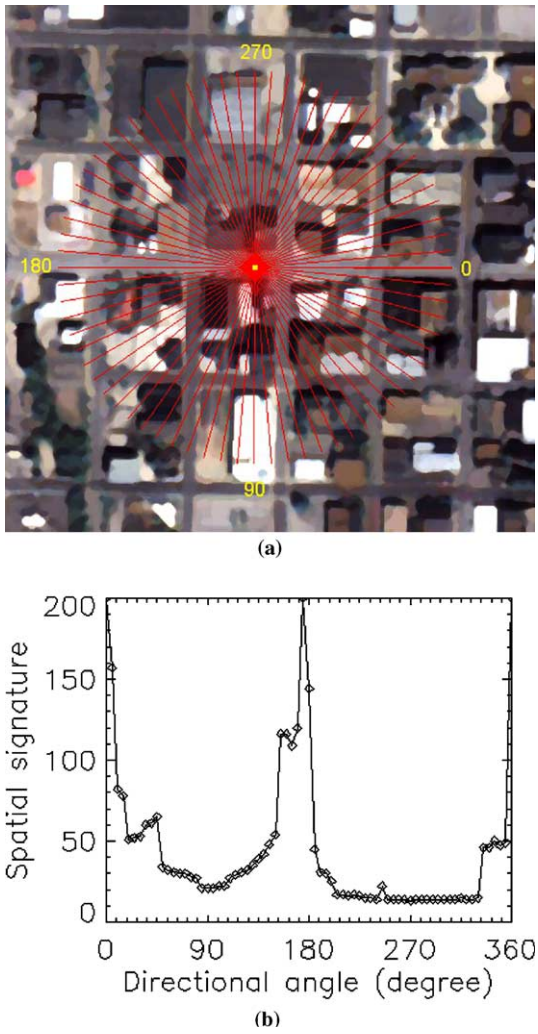

Fig. 2. Illustration showing the spatial signature of a road pixel: (a) searching along multiple radial directions for a road pixel and (b) the corresponding 1D spatial signature of the road pixel.

2.3. Dense urban area identification

We first differentiate dense urban areas from suburban areas to allow the appropriate algorithms to be applied (Fig. 1). Dense urban areas are characterized by low vegetation density and high density of commercial buildings and other impervious surface (e.g. parking lots). Here we assume that a dense urban center covers an area of at least six city blocks along each of two perpendicular directions. For 1 m PS-MS IKONOS imagery, the distance of most city blocks in the test image was empirically observed to be about 100 m or pixels. After scanning the entire image, the 600×600 area with the lowest average NDVI value is found. The density of commercial buildings could be an auxiliary criterion, but this was not used here. After identification

of the dense urban center, a tiling operation was carried out iteratively around the detected urban center. For each 100×100 candidate tile, the average NDVI value was checked. If the candidate tile had an NDVI value similar to the urban center, the candidate tile was added to the urban center and further tiling around it is then explored. Otherwise, the iterative checking of other candidate tiles is stopped. The detected dense urban area from the IKONOS image of Columbia, Missouri is shown in the box surrounded by the highlighted solid line in Fig. 4(a). In the detected area, most roads are a series of parallel or orthogonal straight lines that group together to form a grid structure. Different road extraction and fusion strategies for application to suburban and dense urban areas are presented in the following two sections.

3. Suburban area road extraction and fusion

In suburban areas, roads are modeled as curvilinear and homogeneous regions with nearly constant width. We used two complimentary detectors to identify candidate roads or road "seeds" from the images. The results from the two detectors are then fused based on an optimum path search algorithm using a cost image. A road tracker is then used to connect missing parts of the road network.

3.1. Detector I: segmentation and grouping

The first detector follows a traditional bottom-up image analysis strategy. First the entire image is segmented into individual, non-overlapping homogeneous regions. Candidate road segments are selected from the images based on shape information from each segment. Next, candidate segments are grouped based on perceptual theory [19,28].

There are a wide variety of segmentation techniques discussed in the literature [10,29]. Traditional gray-scale image segmentation is based on edge detection or region growing. In [10], region information and edge information are combined in a global optimization framework by maximizing a fuzzy-based evaluation function. However, this algorithm is computationally intensive for traditional computers, e.g. non-parallel. The success of the image segmentation is critical for higher-level tasks such as feature extraction and object recognition.

With the 1 m PS-MS image resolution and nominal swath width of 11 km, IKONOS satellite images have a large image size. The K-Means clustering method was applied to the masked image, and this particular segmentation approach is computationally fast and effective for high-resolution multi-spectral imagery. The segmented image is then obtained through connected component labeling on the classified image.

Segments are selected as road seeds if the following two criteria are satisfied: (1) the mean width of the segment is within a certain range, (2) the length to width ratio is larger than a preset threshold. In addition, some shadows and/or other mixed-pixel regions may also have a quasi-linear structure, but these usually have a relatively smaller area/size than road segments. Thus, a threshold based on segment area/size is also used as a constraint during road seed selection.

After selection of the road seed segments, we use a modified Wang–Zhang thinning algorithm [13] to extract skeletons from the road segments. While there are many thinning algorithms in the literature [14], some do not preserve the structure of the original objects. We selected the modified Wang–Zhang thinning algorithm because it does not shorten the ends of the segment and its skeleton product is connected. A morphological closing operation is applied before thinning to smooth the segments. After thinning, roads are reduced to their centerline representation with a one-pixel width. A pruning operation is then used to remove short dangling branches of the centerlines caused by, for example, driveways.

Despite the preprocessing steps to smooth out spectral variations, the segmentation results can be highly fragmented due to significant variations in road surface reflectance. As a result, some roads can be broken into several parallel segments. In addition, shadows from trees and buildings cause many road segments to be disconnected. In humans, perceptual organization is the ability to readily group elements in an image based on various relationships between them such as similarity, proximity, colinearity, parallelism, and connectivity [19]. Here, perceptual grouping operations are applied on two scales. Grouping is first applied on the original segments to group redundant or overlapped segments, after which the centerlines are then regenerated. Then, disconnected centerlines are grouped based on proximity and colinearity. The perceptual road-seed grouping is done using the following rules:

(1) If two segments are adjacent and most parts are parallel, then they are combined into a new segment and a new centerline is generated.
(2) If two segments are adjacent at their two ends, and the average fuzzy edge magnitude between them is small, then they are combined into a new segment and a new centerline is generated.
(3) If the endpoints of two separate centerlines are within some small distance, then these lines are connected.
(4) If the endpoints of two centerlines are separated by a moderate distance, but the line segments are in the same direction, then these two lines are connected.

Examples of road segment grouping are shown in Fig. 3. By segmentation and grouping, a large number

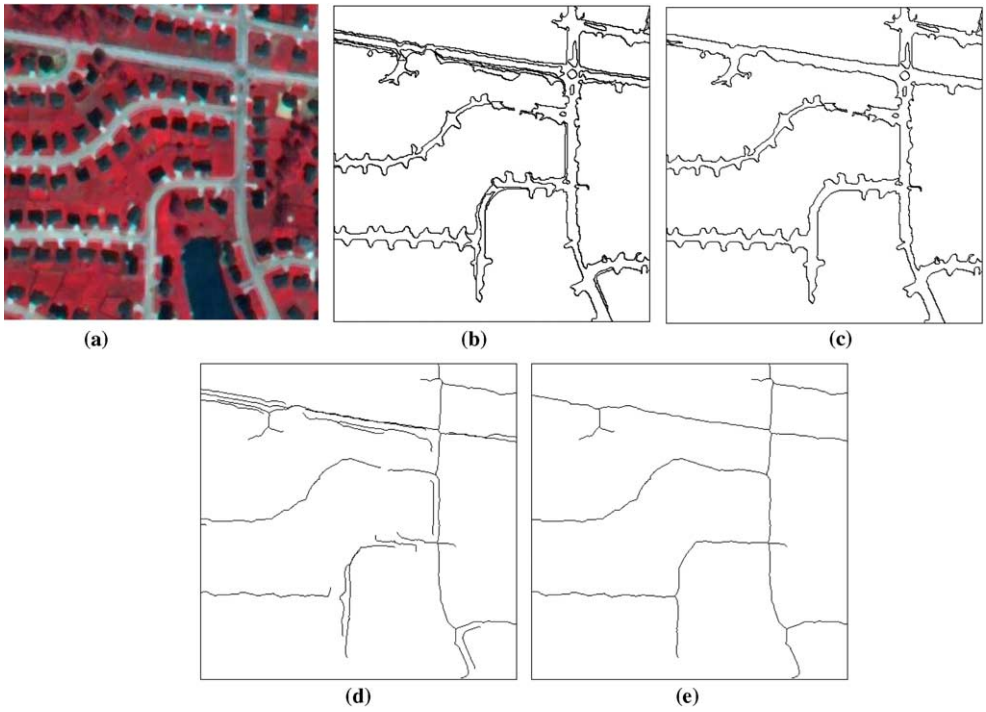

Fig. 3. Perceptual road grouping on two scales: (a) image subset after filtering (NIR, R, G channels are shown); (b) road seed segments before grouping; (c) road seed segments after grouping; (d) centerlines before grouping and (e) centerlines after grouping.

of road segments are detected and the final road centerline selections are generated. The strength of this detector is that all selected centerlines have a very high probability of belonging to the road network. However, a significant number of roads are still not detected in this approach.

3.2. Detector II: curvilinear structure detection

There is significant spectral variability in road reflectance within a given image because of different ages, construction materials, and illumination angles. Some roads will be severely fragmented in the segmentation detector and cannot be detected using their shape information. In addition, some roads are segmented with new construction areas (due to the spectral similarity) which completely alters the elongated characterization of the road segment. Thus, we require an additional road seed detector using another information source: curvilinear feature detection.

Roads can appear as dark or bright lines in an image depending on age and/or composition. One method for line detection is to locate positions of ridges and ravines in the image function and extract lines from these by using differential geometric properties [27]. In this approach, roads are modeled as having 1D bar-shaped or parabola-shaped intensity profiles perpendicular to road direction. The first-order derivative of the line center should vanish and the second-order derivative of the line center should be of large absolute value. However, in practice this is not always true. Sometimes, one side of the road is darker and the other side is lighter. In this situation, the absolute value of the second-order derivative of the line center is zero (not a maximum) and the line cannot be detected [11]. However, as long as the line segment is homogeneous, the line can be detected in the segmentation step. So segmentation and curve detection are complementary approaches.

In suburban areas with a lot of vegetation, roads have low NDVI values and appear as dark lines in the NDVI channel. We extracted ravines from the NDVI channel using a curvilinear structure detector. While curve detection can also be done on the PAN channel, this can sometimes induce false positives and was therefore not adopted here. For curve detection, the NDVI image was first smoothed by convolving it with a Gaussian kernel. In the smoothed coarser image, lines can be extracted even in presence of some disturbances. The 1D Gaussian kernel is given by

$$g_\sigma(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-((x^2)/2\sigma^2)}. \quad (3)$$

The partial derivatives r_x, r_y, r_{xx}, r_{xy}, and r_{yy} of the image are estimated by convolving the image with the discrete 2D Gaussian partial derivative kernels. Then, the local line direction of each pixel is determined by calculating the eigenvalues and eigenvectors of the Hessian matrix

$$H(x,y) = \begin{pmatrix} r_{xx} & r_{xy} \\ r_{xy} & r_{yy} \end{pmatrix}. \quad (4)$$

The response of the line detector at each pixel can be determined by the first-order and second-order derivatives along the normal direction of the line. If the first-order derivative along the normal direction vanishes within the current pixel, the second-order derivative along the normal direction is taken as the response of the line detector at the current pixel, otherwise the response is set to zero.

We used the curvilinear structure detector on multiple scales by choosing Gaussian kernels with different smoothing parameters σ [15]. Road centerlines have large response on multiple scales, whereas other features may have a large response on only one scale. The line-detector response from multiple scales was obtained by summing responses from all scales and then transforming this to a fuzzy decision of road centerlines using an S-type function [10]. In the fuzzy membership, a low threshold is set to zero and a high threshold is adopted so that all pixels with response larger than the high threshold are set to 1.0. The intermediate fuzzy values between 0 and 1 are used later to construct a cost image for fusion of the results from the two complimentary detectors.

3.3. Detector fusion results

3.3.1. Generation of additional road seeds from detector II

By segmentation and grouping, candidate road centerlines were detected and hard decisions were made. However, a significant number of road segments were not identified by this detector. We expect that additional road seeds from the complimentary second detector can be used to improve the overall results.

From the curvilinear structure detector, a membership value from 0 to 1 was obtained for each pixel representing the likelihood that it belongs to a road centerline. Many pixels have a very high membership value. Consequently, a ribbon-shaped mask is constructed along the road centerlines identified by the segmentation detector. The pixels outside the mask that have a membership value larger than a threshold are then recorded. The connected lines constructed by the recorded pixels that are longer than a preset value are considered as additional road seeds. The additional road seeds are combined with the initial road seed group using an "OR" operation.

3.3.2. Cost image construction

From detector II we have a membership value for each pixel that represents the degree that the pixel belongs to the road network. The range of the fuzzy membership value is $[0,1]$. If the membership value is high, then the cost to pass it into the selected road set is low; otherwise the cost is high. We use a simple mathematical operation to transform the fuzzy values into cost values

$$C = 1.0 - f, \quad (5)$$

where C represents the cost value of each pixel, and f represents the fuzzy membership value generated by curvilinear structure detector.

For those pixels already belonging to the road seed set from detectors I and II, the cost value to pass them is set to a very low value, such as 0.05. It cannot be set to zero because most search paths will try to pass through the detected road seeds regardless of their length. For those pixels identified in areas of vegetation and water bodies, we set the cost value to a high value (e.g. 1.0) to ensure that the search path avoids these areas.

3.3.3. Fusion by optimum search path

From a global perspective, roads are an interconnected network of linear and/or curvilinear lines and are seldom separate pieces. In the preceding sections, we generated road seeds based on two different detection strategies. The road seeds generated from the two different sources are seldom connected because of the relatively low membership values of the pixels between them. In this step, we try to connect the additional road seeds from the second detector to the original road seeds through optimum path searching over the cost image.

First, the nodes of the original road seeds are detected. The nodes include endpoints, intersections, and corners. Corner detection is based on the curvature at each pixel and here we adopt the k-curvature described in [17]. Curvatures for pixels along the line are then smoothed by convolution with a Gaussian kernel. A corner is detected based on three conditions: (1) the curvature at the corner is larger than a threshold, (2) the curvature at the corner is a local maximum, (3) the pixel at the corner has a large displacement away from a B-spline passing through the surrounding pixels on the line.

After identifying the nodes of the original road centerlines, the endpoints and their extension directions of the additional road seeds are detected. For each endpoint of a new detected line, the nodes within a small distance and a certain angle range are searched. An attempt is then made to connect the endpoint to surrounding nodes using the shortest path search on the cost image. Dijkstra's algorithm [3] is widely used for this purpose because of its high computational efficiency. For each optimum path connecting the endpoint of the additional road seeds and the nodes of the original road seeds, the average cost of the path is evaluated. If it is lower than a threshold, then this connection is considered valid and all the pixels along the search path are set as road pixels.

By fusing the results of the two detectors, most parts of the road network can be detected. However, some

parts of the road network are missed because of their low contrast or high spectral variation. Using the results from the two automated detectors, a large number of reliable road vectors and their directions are now known. These can then be used in a road tracker to follow the missing parts of the road network.

3.4. Road tracking

Here the road centerlines detected by fusing the two detection results are tracked at each node to extrapolate missing parts of the road network. A road tracking strategy by profile matching [32] is adopted here. First, a typical intensity profile perpendicular to the road axis is extracted and used as a reference profile. The next position of the road axis is predicted through linear extrapolation to several pixels beyond the road centerline. Profiles near the new position (search profiles) are matched with the reference profile. The best matching position is determined by the least squares correlation criterion. Here we use the intensity profiles of the NDVI *and* PAN channels. The least squares correlation criterion minimizes

$$\sum_{i=1}^{n}\sum_{j=1}^{n_b}(r_{ij}-s_{ij})^2, \tag{6}$$

where r_{ij} is the intensity of the jth channel of the ith pixel on the reference profile, s_{ij} is the intensity of the jth channel of the ith pixel on the search profile, n is the profile length, and n_b is the number of channels (two in this application).

The next road centerline point is estimated by weighting the predicted position and the best profile matching position. In experiments, we found that the matching step is more reliable, so a high weight (0.7) is set for the best matching position. After tracing one step ahead, the reference profile is updated by weighting the original reference profile and the best matching search profile. A low weight is set for the best matching search profile to keep the spectrum of the tracked roads similar to their parent road pieces. The tracking continues until an already extracted road is reached or multiple poor matches are encountered. If continuous pieces of the tracked roads belong to forbidden areas such as vegetation, the tracking also stops.

4. Urban area road extraction and fusion with suburban area results

In practice, dense urban areas cannot be totally separated from surrounding suburban areas. In this section, we describe the technique used to extract the roads in urban areas and then fuse the urban-area results with those from the surrounding suburban areas.

4.1. Urban area road extraction

For many US cities, roads often appear as straight lines that form a grid structure in dense urban areas. A line segment matching method has been proposed for extracting urban road networks from high-resolution satellite images [24]. However, this approach assumes the outline of the road network is readily visible on a binary image produced by simple thresholding applied to the original gray scale image. The line segments are then matched with the binary images in all directions to detect all the possible components. The Hough transform is a well-known method for line detection from raster images. Dell'Acqua and Gamba [1] used two modified versions of the Hough transform to detect roads from high-resolution SAR images. First they used a fuzzy C means (FCM) algorithm to classify the image. Each pixel was then assigned a street membership value. Then the connectivity weighted Hough transform (CWHT) and the rotation Hough transform were modified to work with fuzzy membership values [1,2]. In this paper, we develop a spatial signature weighted Hough transform (SSWHT) that takes the spatial signature (Fig. 2) of each pixel into account. An added benefit is this modified Hough transform can be applied to multi-spectral images.

4.1.1. Removing short-scale structures

In dense urban areas, there are often many dark and/or bright structures such as buildings or parking lots. These features are usually shorter than the distance of a city block. However, if many features appear along on a line, there will be a big response in the polar plane (Hough transform). Here we adopt a directional morphological filter [26] to mask these short-scale structures out.

The directional morphological filter is applied to the PAN image. From empirical observations of the 1 m IKONOS image, a city block distance is about 100 meters or pixels. This parameter can be set by the user for different circumstances. Bresenham lines of a length a little larger than the city block distance, such as 120 m, are used for the structure elements (SEs). We adopted an operation called union of the openings. First, a series of directional morphological openings with line SEs of different directions are applied to the image. The union of the openings is then obtained by taking the maximum values of all openings. Usually the road pixels have a length larger than 120 m and have an inconsistent contrast with the background. In the union of the openings image, the road intensity is very stable, whereas the short bright structures will vanish by the union of directional openings. A top-hat operation is executed by subtracting the image of the union of the openings from the original panchromatic image. We call this a directional morphological top-hat operation. Similarly, a directional morphological bottom-hat operation is defined as

subtracting the original panchromatic image from the image of the intersection of the closings. By setting a threshold on the directional top-hat images, bright structures shorter than a city block can be masked out. Similarly, dark structures shorter than a city block can be masked out by setting a threshold on the directional bottom-hat images. Fig. 4(a) shows the panchromatic image of the urban–suburban test site. The results of the directional top-hat and bottom-hat operation are shown in Figs. 4(b) and (c), respectively. The short structure mask is shown in Fig. 4(d).

4.1.2. Road grid hypothesis by SSWHT

In the original Hough transform, a straight line can be represented by the polar equation

$$x\cos\theta + y\sin\theta = \rho, \tag{7}$$

where θ and ρ are line parameters. Each pixel (x_i, y_i) in the original image plane corresponds to a curve in the (ρ, θ) plane. All pixels on the same line intersect at a point in the (ρ, θ) plane. The parameters θ and ρ of the intersection point in (ρ, θ) will define the line. If we count the number of curves passing at each point on the (ρ, θ) plane, the lines in the original image plane can be detected by identifying the points in the (ρ, θ) plane that have the highest accumulated values. The points are then backprojected to the original image plane to represent the detected line.

After masking out bright or dark structures shorter than a city block, the road grids are readily seen. However, there are still many disturbances. For example, some parking lots are connected to the road network because they do not have high contrast with the surrounding background. Thus, they cannot be masked out by directional morphological filtering and aligned structures will generate a high response in the polar plan. The normal Hough transform will not work well on the original image. Here we use information obtained from the spatial signature measure described in Section 2.2 to enhance the roads.

For the spatial signature of each pixel, the length is recorded as the maximum length of a homogeneous line

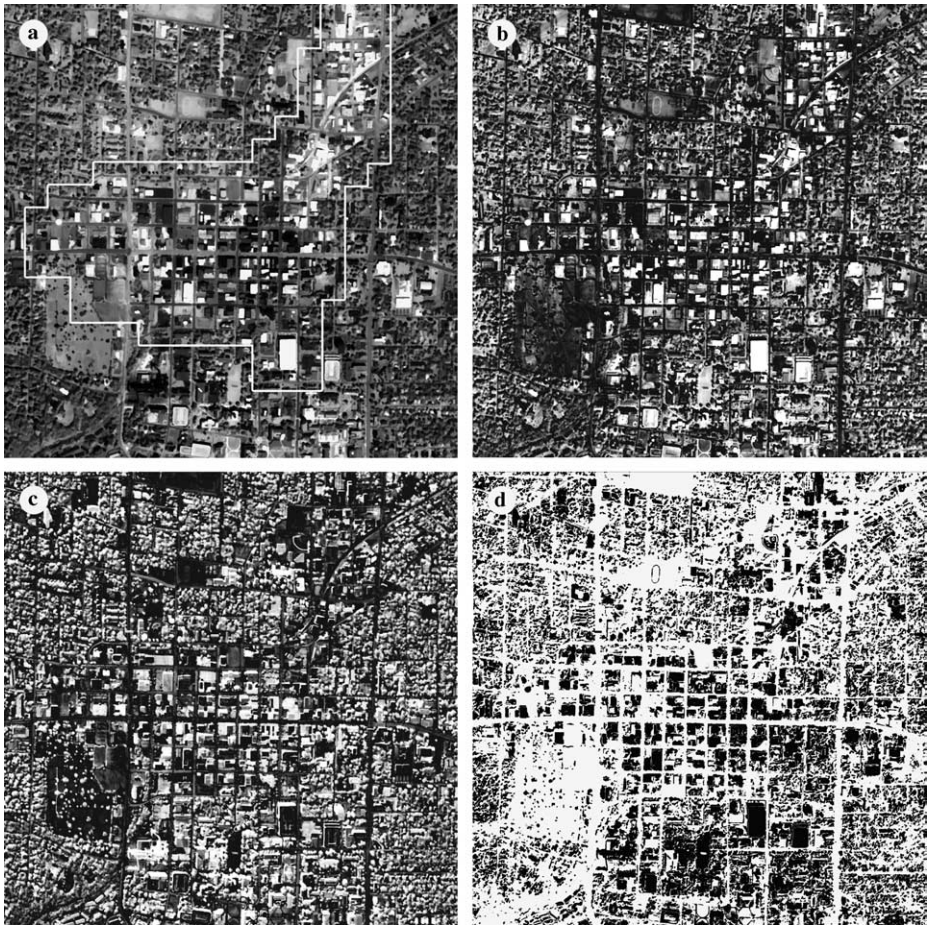

Fig. 4. Example of directional morphological filtering applied to high-resolution satellite image: (a) panchromatic image of the urban–suburban test site (the identified dense urban area is surrounded by the highlighted white solid line); (b) results of the directional top-hat operation; (c) results of the directional bottom-hat operation and (d) structures with a short spatial scale (<1 city block) are masked out (black).

passing through the current pixel. Because roads are usually homogeneous straight lines in urban areas, road pixels will generally have a much larger length measure than the pixels of other shorter-scale features. Here we use the length measure from the spatial signature as a weight for the Hough transform. We call this the spatial signature weighted Hough transform (SSWHT). For each pixel (x_i, y_i) in the original image plane, we add its length measure rather than 1 to the accumulated response at the points on the corresponding curve in the (ρ, θ) plane. In this way, the point in the (ρ, θ) plane with the highest accumulated response will correspond to the road with the longest length in the original image plane.

In digital images, exact co-intersection at one point by many (ρ, θ) curves may not be possible because of quantization. In addition, roads occupy a certain width, and regions with similar values of θ and ρ will have high responses rather than a point. Accordingly, the road grid hypothesis is generated using the following steps:

(1) In the polar plane, the point (ρ_0, θ_0) with the highest response is determined. The θ_0 line parameter is used as an estimate of one of the main directions of the road grid.
(2) The responses in a small angle range $[\theta_0 - \Delta\theta, \theta_0 + \Delta\theta]$ for all ρ are accumulated. The histogram of responses vs. ρ are then plotted.
(3) The ρ_i, $i = 1, 2, \ldots, k$, corresponding to the salient peaks in the histogram are found. This is implemented on two scales: First, peaks with a magnitude larger than a threshold are found and then the average distance between the peaks is taken as an estimate of the city block distance. Then, peaks with a magnitude larger than a lower bound that are also at least a distance of three fourths of a city block away from any peaks detected in the first step are found.
(4) For each position ρ_i, $i = 1, 2, \ldots, k$, the direction θ_i is chosen in the range from $[\theta_0 - \Delta\theta, \theta_0 + \Delta\theta]$ which have the maximum response at the position ρ_i.
(5) Usually another direction is perpendicular to the first direction identified in Step 1. Accordingly, $\theta_1 = (\theta_0 + 90) \bmod 180$ is used as an estimate of the other main direction. The same method described in Steps 2–4 is used to find the exact directions and positions of the centerlines along the second direction.
(6) After detecting lines oriented along the two main directions, regions around the two directions are masked out in the (ρ, θ) plane. Remaining points with high response can be detected if any are left.
(7) The points (ρ_i, θ_i) are backprojected to the original image to determine the hypothesized road grid. For the urban–suburban test site examined here, the hypothesized road grid is shown in Fig. 5(c).

4.1.3. Road grid verification

After generating the road grid hypothesis, the nodes of the grid are then found. Each centerline piece between two adjacent nodes is verified individually on the original image. First, a ribbon with a width approximately equal to the road width is built around the centerline piece to be verified. The pixels in this ribbon are checked. If all of the following conditions are satisfied, then this centerline is considered valid:

(1) The pixels in the ribbon are homogeneous. This is measured by the standard deviation of the pixel intensities. The standard deviation should be less than a preset threshold.
(2) Most of the pixels are in valid regions, i.e. they do not belong to vegetation or short bright/dark structures.
(3) Some of the pixels have a spatial signature longer than the city block distance and the width is within the range of road widths (8–30 m).

The verified road grid based on the above criteria for the urban test site is shown in Fig. 5(d). In the image, the hypothesized and verified road grid are in very good agreement with the reference road network shown in Fig. 5(b) developed by manual extraction.

4.2. Fusion with suburban area results

Dense urban areas are usually surrounded by suburban areas. There are often transitional areas between these two environments. After detecting the road networks from the two environments separately, the results need to be fused together.

In Section 2.3, we described the technique used to identify dense urban areas. A dilation operation with a disc-shaped SE of radius $r = \textit{city block distance}$ is applied to the dense urban area to identify transitional areas. First, the road grid hypothesis and verification is implemented on both the dense urban areas *and* the transitional areas. At the same time, road extraction by fusion of the two complementary detectors is done on the suburban *and* the transitional areas. In the suburban areas, we use the extracted centerlines generated from the techniques described in Sections 3.1–3.3. In the urban areas, we use the results from the approach described in Section 4.1. In the transitional area, we fuse the two results together. Here we treat the results from the road grid hypothesis and verification as a more reliable source because it is based on a global perspective. We accept all the verified centerlines as road centerlines. A ribbon is used to mask out areas around the verified centerlines. The centerlines detected by fusion of the two detectors are checked. If these centerlines reside in the masked area and are parallel to the city main directions, then they are deleted. Other centerlines are

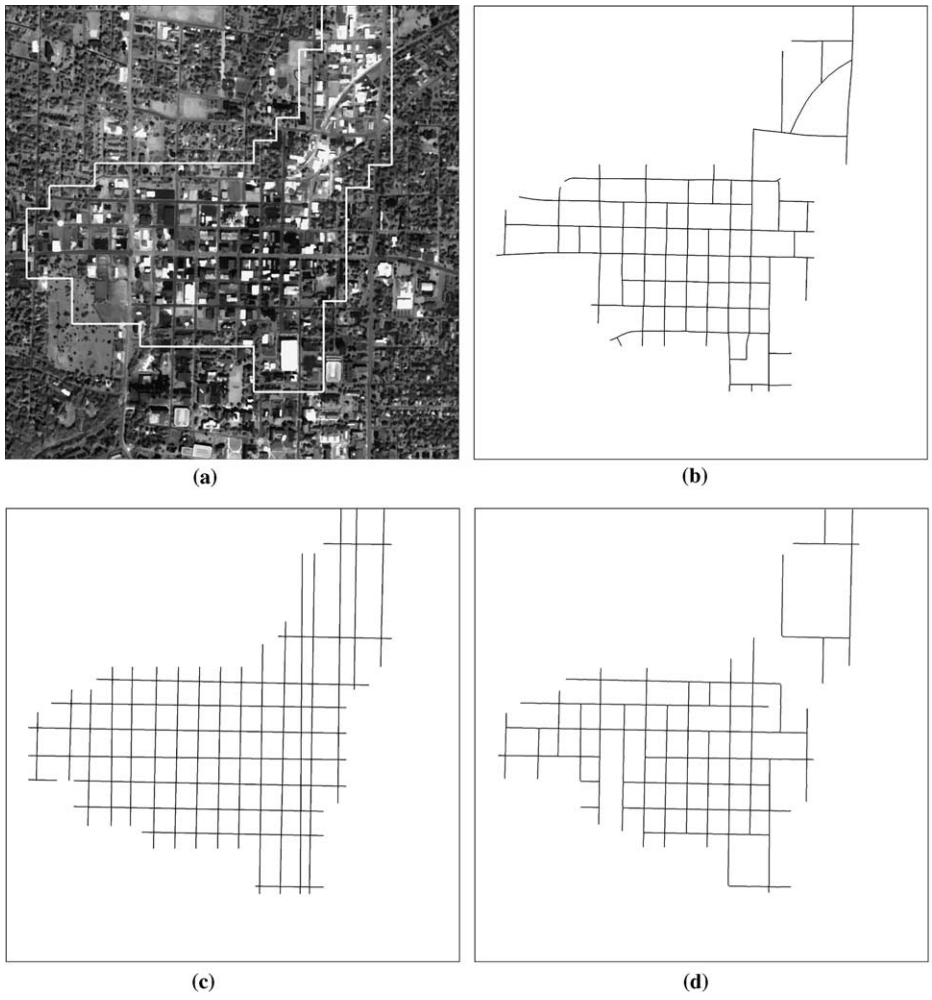

Fig. 5. Road extraction on dense urban area: (a) panchromatic image of the urban–suburban test site (the identified dense urban center is surrounded by the white solid line); (b) reference road network; (c) road grid hypothesis by SSWHT and (d) verified road grid.

kept and combined with the road grid. After detecting and fusing the results on the urban and suburban areas, the centerlines are used as starting points for the road tracker described in Section 3.4 to extrapolate missing pieces of the road network.

5. Results and evaluation

The integrated road extraction system based on multi-detector information fusion is tested using the 1 m PS-MS IKONOS image of the City of Columbia, Missouri acquired on April 30, 2000.

5.1. Study area

The imagery used for this study was acquired by the IKONOS commercial remote sensing satellite and consists of four multi-spectral (MS) bands with 4 m resolution and a single panchromatic (PAN) band with 1 m resolution. Both the PAN and MS data have 11-bit information content. The 4 m MS data consists of four individual bands: Red (R), Green (G), Blue (B), and Near Infrared (NIR). The test image includes a variety of urban and suburban land cover types that make it ideal for this study.

Three representative test sites with a size about 4 km^2 were selected from the imagery. The first two test sites were selected from suburban areas. They include most of the complex situations encountered in suburban areas that make road extraction difficult: dark vs. bright roads, roads vs. new construction areas, roads vs. creeks, roads vs. bare soil, etc. The third test site was selected in an urban–suburban area. This test site was used to validate the road extraction strategy on the urban area and the fusion strategy by combing the urban and suburban road extraction results.

5.2. Evaluation criteria

The extracted centerlines were compared with manually plotted vector data (reference data). Quality

measures proposed by Wiedemann et al. [34] were adopted to evaluate the quality of road extraction. These measures are defined as follows:

$$\text{Completeness} = \frac{\text{length of matched reference}}{\text{length of reference}} \tag{8}$$

$$\text{Correctness} = \frac{\text{length of matched extraction}}{\text{length of extraction}} \tag{9}$$

$$\text{Quality} = \frac{\text{length of matched extraction}}{\text{length of extraction} + \text{length of unmatched reference}} \tag{10}$$

In the evaluation, a buffer with a width equal to the road width is built along the reference data and extracted centerlines. If a pixel on the extracted centerlines is within the reference data buffer, then it is counted as a matched extraction. If a pixel on the reference vector data is within the buffer of the extracted centerlines, it is counted as a matched reference.

5.3. Suburban test site results

The road extraction steps described in Section 3 were applied to the two suburban test sites. First, the initial road seeds were selected by segmentation and grouping. Then, additional road seeds were generated by the curvilinear structure detector and fused with the initial road seeds. Finally, a road tracker was used to fill in missing pieces of the road network. The evaluation results after each key processing step are summarized in Tables 1 and 2. The road centerline extraction results from test sites I and II are shown in Figs. 6 and 7, respectively.

Table 1
Evaluation results of suburban test site I

	Completeness	Correctness	Quality
Detector I	0.81	0.91	0.75
Fusion of two detectors	0.84	0.91	0.78
Road tracking	0.86	0.92	0.80

Table 2
Evaluation results of suburban test site II

	Completeness	Correctness	Quality
Detector I	0.70	0.69	0.54
Fusion of two detectors	0.72	0.69	0.55
Road tracking	0.78	0.70	0.58

From the tables, the effectiveness of each step in the integrated extraction strategy is shown. It is not surprising that after each step, the completeness increases as missing roads are identified and connected. In the road completion steps, some constraints such as forbidden areas are defined so that the correctness stays at about the same level although more pieces are added to the road network. As a whole, the quality measure increases after each step.

For the two suburban test sites, the completeness ranges between 78% and 86% and the correctness ranges between 70% and 92%. These results are among the highest levels reported in the existing literature for suburban areas [7,8]. However, the correctness for test site I is significantly higher than test site II because roads are confused with building roofs and bare soil boundaries in test site II during the road seed selection using shape information. Thus, it is clear that the first

Fig. 6. Road extraction results from suburban test site I. Note the spectral variations in the road surfaces and the proximity of some roads to new construction areas: (a) original IKONOS image; (b) manually plotted road centerlines (reference) and (c) road centerline extraction results.

Fig. 7. Road extraction results from suburban test site II: (a) original IKONOS image; (b) manually plotted centerlines (reference) and (c) road centerline extraction results.

step—road seed selection—is critical to overall success of the algorithm. We prefer omission errors to commission errors in this step to limit false positives with the hope that the line detection and road tracking can be used to improve the road network after road seed selection.

5.4. Urban–suburban test site results

This test site was selected to cover the urban center of the test image. The dense urban area was automatically identified as described in Section 2.3 and it is shown in Fig. 5(a). To emphasize the special characteristics of urban areas, we first present the extraction results on this test site using only the suburban strategies. The extraction results are shown in Fig. 8(c). From the results we can see that the suburban strategies failed many times in the urban area where roads are routinely connected with surrounding compact structures such as buildings or parking lots. Next, the extraction results by fusing urban and suburban results are reported. Finally, the road network was completed by a road tracker. A comparison of the evaluation results is provided in Table 3. The final extraction results by information fusion are shown in Fig. 8(d).

From the table we see that the extraction results are greatly improved by using different strategies for the dense urban area. In the final results, the completeness of the extracted roads is 70% and the correctness is 79% for the complex urban–suburban area. Seldom have road-extraction results from high-resolution satellite images been reported in the literature for areas such as this. The most related work is that of Haverkamp et al. [8], where 49–66% road extraction completeness was achieved for a set of four IKONOS image subsets (each 2048×2048 pixels) mixing urban and suburban content. Here, our extracted road network achieves significantly better results with completeness values that range between 70% and 86% and correctness values that range between 70% and 92% over all test sites.

6. Discussion and conclusion

In this paper we presented an integrated system for urban and suburban road mapping from high-resolution satellite imagery using multi-detector information fusion. This system incorporates multiple automated road detectors and a road tracker. The road detectors are chosen differently for different environments. In dense urban areas, roads usually appear as a series of parallel and orthogonal straight lines that form a grid structure. We proposed a SSWHT algorithm to hypothesize the road grid from a global perspective. The urban area road extraction was then based on street grid hypothesis and verification. In suburban areas, roads are modeled as curvilinear and homogeneous segments with nearly constant width. Candidate road seeds were obtained by fusing the results of two complementary road detectors. The road centerlines detected in suburban and urban areas were fused together and a road tracker follows to complete the road network. The integrated system is implemented automatically without the need for manual selection of road points.

Based on the experimental results presented in the previous section, we conclude that road extraction results can be greatly improved by fusion of multiple road detectors developed from different models that are then automatically applied to different contextual environments. By fusion of multi-detector information, our integrated road extraction system obtained very good results for complex urban–suburban environments. The results from three test sites for suburban and urban areas were reported. The road extraction completeness ranged between 70% and 86% and correctness ranged

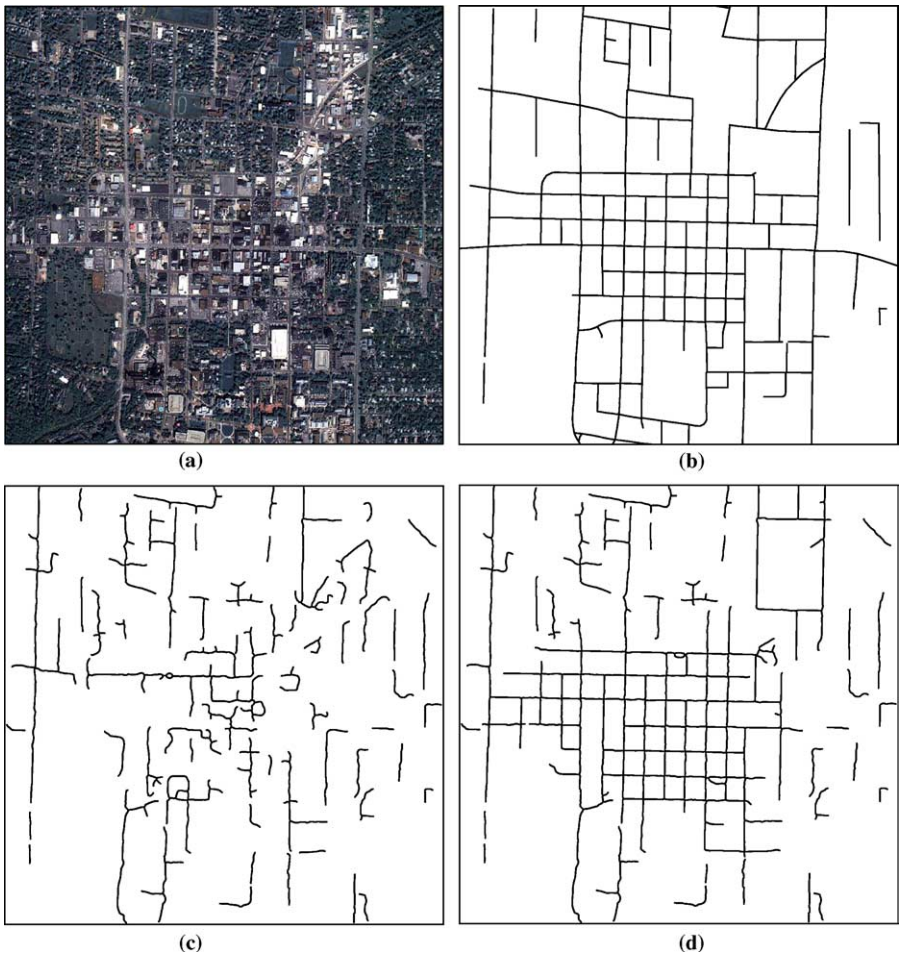

Fig. 8. Road extraction results from urban–suburban test site: (a) original IKONOS image; (b) manually plotted centerlines (reference); (c) road centerline extraction results using only suburban strategies and (d) road centerline extraction results by fusing suburban and urban strategies.

Table 3
Evaluation results of urban–suburban test site

	Completeness	Correctness	Quality
Suburban strategies	0.46	0.67	0.37
Detector fusion and SSWHT	0.65	0.81	0.56
Road tracking	0.70	0.79	0.59

between 70% and 92%. These extraction results are among the highest reported thus far in the literature.

For urban areas, the spatial signature weighted Hough transform (SSWHT) was used to extract linear road structures for relatively regular urban street grid patterns. This approach can be easily adapted to extract linear structures along any direction as needed. The SSWHT has unique advantages because it incorporates local shape information into the linear feature analysis and can be applied to multi-spectral imagery. The urban grid model can be applied to many US cities. However, in some cities the model may not be appropriate, e.g. some towns in hilly environments, some older European cities, etc. Other models and road extraction algorithms would need to be developed and then incorporated into the integrated system to handle these situations. The key issue is the discrimination of different contextual environments to allow an integrated system to adaptively select the appropriate road detectors for the specific contextual setting. This paper demonstrates this adaptive approach for a typical US city. The integrated system is an open architecture and can be enriched and improved further by fusing with other extraction methods as needed.

In this paper, the integrated system for automatic road mapping was applied to high-resolution multispectral satellite imagery. Key algorithm parameters reported in this paper are appropriate for IKONOS PS-MS data with 1-m resolution. The system can be extended to other high-resolution satellite imagery (e.g. QuickBird) by changing the metric distances into pixel units according to image scale. For aerial imagery with similar image resolution, the same extraction strategies can also be easily adapted. The SSWHT will work on color imagery without modification. However, for

vegetation discrimination some greenness measure must be used in place of the NDVI measure used here. The integrated system architecture outlined here can be extended to road extraction on imagery with other image resolutions as well. However, for imagery with different resolutions, the difference in object scale and the details discriminated in the imagery will vary dramatically. For example, for aerial imagery with resolution less than 0.3 m, vehicles can be more easily identified in the imagery and provide supporting contextual information for the road extraction. Thus, other models and algorithms appropriate for even higher image resolutions than that utilized here can be incorporated into the system as well. The objective of this research was to demonstrate that multiple road detector strategies could be adaptively applied to high-resolution satellite imagery based on the contextual environments and then followed by fusion for final road network generation.

From the extraction results we can see that most errors occur due to disturbances from surrounding objects. For examples, trees fully occlude some road segments in suburban test site I and the urban test site. In addition, some building roofs, bare-soil boundaries, and linear shadows were erroneously extracted as roads in suburban test site II. Thus, the road extraction is clearly influenced by surrounding objects. However, surrounding objects can sometimes provide supporting contextual information for the road extraction. For example, in suburban areas some tree lines and curvilinear groups of single-family houses are parallel to the road. In urban areas most building boundaries are parallel to the road grid, and vehicles detected on the road can also help to verify some road segments. We are presently working on the automated extraction of other urban-related features so that mutual verification strategies can be applied to improve the extraction of multiple features such as roads, vehicles, and buildings.

Acknowledgements

This research was supported by the Raytheon Synergy program under subcontract #3000615 from NASA.

References

[1] F. Dell'Acqua, P. Gamba, Detection of urban structures in SAR images by robust fuzzy clustering algorithms: the example of street tracking, IEEE Trans. Geosci. Remote Sens. 39 (10) (2001) 2287–2297.
[2] F. Dell'Acqua, P. Gamba, G. Lisini, Road map extraction by multiple detectors in fine spatial resolution SAR data, Can. J. Remote Sens. 29 (4) (2003) 481–490.
[3] E.W. Dijkstra, A note on two problems in connection with graphs, Numer. Math. 1 (1959) 269–271.
[4] M.A. Fischler, J.M. Tenenbaum, H.C. Wolf, Detection of roads and linear structures in low resolution aerial images using multi-source knowledge integration techniques, Comput. Graph. Image Process. 15 (3) (1981) 201–223.
[5] P. Gamba, O. Hellwich, P. Lombardo, Theme issue on algorithms and techniques for multi-source data fusion in urban areas, ISPRS J. Photogrammet. Remote Sens. 58 (1–2) (2003) 1–3.
[6] R.C. Gonzalez, R.E. Woods, Digital Image Processing, second ed., Prentice-Hall, NJ, 2002.
[7] B. Guindon, A framework for the development and assessment of object recognition modules from high-resolution satellite images, Can. J. Remote Sens. 26 (4) (2000) 334–348.
[8] D. Haverkamp, R. Poulsen, Complementary methods for extracting road centerlines from IKONOS imagery, in: Proceedings of SPIE Vol. 4885, Image and Signal Processing for Remote Sensing VIII, SPIE, 2002, pp. 501–511.
[9] S. Hinz, A. Baumgartner, Automatic extraction of urban road networks from multi-view aerial imagery, ISPRS J. Photogrammet. Remote Sens. 58 (1–2) (2003) 83–98.
[10] X. Jin, C.H. Davis, A genetic image segmentation algorithm with a fuzzy-based evaluation function, in: Proceedings of IEEE Conference on Fuzzy Systems, St. Louis, MO, 2003, pp. 938–943.
[11] X. Jin, C.H. Davis, Automatic road extraction from high-resolution multispectral IKONOS imagery, in: Proceedings of International Geoscience and Remote Sensing Symposium, IGARSS2003, Toulouse, France, 2003, pp. 1730–1732.
[12] G.J. Klir, B. Yuan, Fuzzy Sets and Fuzzy Logic: Theory and Applications, Prentice-Hall, NJ, 1995.
[13] R. Krishnapuram, L.F. Chen, Implementation of parallel thinning algorithms using recurrent neural networks, IEEE Trans. Neural Networks 4 (1) (1993) 142–147.
[14] L. Lam, S.W. Lee, C.Y. Suen, Thinning methodologies—a comprehensive survey, IEEE Trans. Pattern Anal. Machine Intell. 14 (9) (1992) 869–885.
[15] H. Mayer, I. Laptev, A. Baumgartner, C. Steger, Automatic road extraction based on multi-scale modeling, context, and snakes, Int. Arch. Photogrammet. Remote Sens. (1997) 106–113.
[16] D.M. McKeown, J.L. Denlinger, Cooperative methods for road tracking in aerial imagery, in: Computer Vision and Pattern Recognition, CVPR'88, IEEE Press, 1988, pp. 662–672.
[17] G. Medioni, Y. Yasumoto, Corner detection and curve representation using cubic B-splines, Comput. Vision Graphics Image Process. 39 (1) (1987) 267–278.
[18] J.B. Mena, State of the art on automatic road extraction for GIS update: a novel classification, Pattern Recognition Lett. 24 (16) (2003) 3037–3058.
[19] R. Mohan, R. Nevatia, Perceptual organization for scene segmentation and description, IEEE Trans. Pattern Anal. Machine Intell. 14 (6) (1992) 616–635.
[20] R. Nevatia, K.R. Babu, Linear feature extraction and description, Comput. Graphics Image Process. 13 (1980) 257–269.
[21] K. Price, Urban street grid description and verification, in: Proceedings of IEEE Workshop on Applications of Computer Vision, 2000, pp. 148–154.
[22] L. Quam, Road tracking and anomaly detection in aerial imagery, in: Proc. of Image Understanding Workshop, 1978, pp. 51–55.
[23] A.K. Shackelford, C.H. Davis, A hierarchical fuzzy classification approach for high-resolution multispectral data over urban areas, IEEE Trans. Geosci. Remote Sens. 41 (9) (2003) 1920–1932.
[24] W. Shi, C. Zhu, The line segment match method for extracting road network from high-resolution satellite images, IEEE Trans. Geosci. Remote Sens. 40 (2) (2002) 511–514.
[25] P. Soille, M. Pesaresi, Advances in mathematical morphology applied to geoscience and remote sensing, IEEE Trans. Geosci. Remote Sens. 40 (9) (2002) 2042–2055.
[26] P. Soille, H. Talbot, Directional morphological filtering, IEEE Trans. Pattern Anal. Machine Intell. 23 (11) (2001) 1313–1329.

[27] C. Steger, An unbiased detector of curvilinear structures, IEEE Trans. Pattern Anal. Machine Intell. 20 (2) (1998) 113–125.
[28] C. Steger, H. Mayer, B. Radig, The role of grouping for road extraction, in: Automatic Extraction of Man-Made Objects from Aerial and Space Images (II), Birkhuser Verlag, Basel, Switzerland, 1997, pp. 245–256.
[29] J.C. Tilton, Image segmentation by region growing and spectral clustering with a natural convergence criterion, in: Proceedings of the International Geoscience and Remote Sensing Symposium, Seattle, WA, 1998, pp. 1766–1768.
[30] F. Tupin, H. Maitre, Detection of linear features in SAR images: application to road network extraction, IEEE Trans. Geosci. Remote Sensing 36 (2) (1998) 434–453.
[31] L. Vincent, Morphological gray scale reconstruction in image analysis: Applications and efficient algorithms, IEEE Trans. Image Process. 2 (2) (1993) 176–201.
[32] G. Vosselman, J. Knecht, Road tracking by profile matching and Kalman filtering, in: Automatic Extraction of Man-Made Objects from Aerial and Space Image, Basel, Switzerland, 1995, pp. 265–274.
[33] J. Vrabel, Multispectral imagery advanced band sharpening study, Photogrammet. Eng. Remote Sens. 66 (1) (2000) 73–79.
[34] C. Wiedemann, C. Heipke, H. Mayer, Empirical evaluation of automatically extracted road axes, In: CVPR Workshop on Empirical Evaluation Methods in Computer Vision, Los Alamitos, California, 1998, pp. 172–187.