A Level Set Method for the Extraction of Roads from Multispectral Imagery

Trish Keaton
Information Sciences Laboratory
HRL Laboratories, LLC *
3011 Malibu Canyon Road
Malibu, CA 90265, USA
pakeaton@hrl.com

Jeffrey Brokish
Electrical & Computer Engineering Dept.
University of Illinois, Urbana-Champaign
1406 West Green Street
Urbana, IL 61801, USA
brokish@ifp.uiuc.edu

Abstract

With the advances in remote sensing technologies, the extraction of roads and other linear features from satellite and aerial imagery has gained substantial interest in recent years. The introduction of satellite imagery characterized by high spectral and spatial resolutions has made possible the development of new viable approaches for the accurate, and cost-effective extraction of linear features with minimal human intervention. This paper presents a semi-automated method for the extraction of roads from high resolution (1-meter) pan-sharpened multispectral IKONOS imagery. An operator provides an initial seed point on the road of interest, then the region is grown using a level set method. Further analysis through iterative smoothing refines the extracted region to accurately estimate the road centerline despite the presence of cars on the road, changes in the pavement or surface properties of the road, or obstruction resulting from foliage or shadows cast on the road by neighboring trees. Initial results have demonstrated the utility of the algorithm in efficiently extracting roads from high resolution satellite imagery with minimal human interaction. Over 97% delineation accuracy was achieved on manually ground truthed IKONOS image samples overlooking both urban and rural locations.

1. Introduction

In recent years, many algorithms have been proposed to extract road networks from satellite imagery. Most of these algorithms [1, 2, 5], however, were developed to extract road networks from panchromatic low resolution satellite imagery (10 meters to 20 meters per pixel) such as SPOT satellite imagery. As more and more high resolution satellite images such as IKONOS images (one meter per pixel) become publicly available, future geographic information systems must support the exploitation of high resolution satellite imagery. The extraction of road networks from high resolution satellite images requires a very different methodology from that used to extract roads from low resolution satellite images. In low resolution satellite images, the width of roads typically ranges from one pixel to five pixels and extraction of road networks is equivalent to detection of lines or detection of strong curvilinear structures [2, 5]. However, with high resolution images, the road-width can very considerably, and additional variations are present such as lane markings, vehicles, shadows cast by buildings and trees, overpasses and changes in surface material. These variations make the extraction of road networks a much more difficult problem.

Recent methods for the extraction of roads from high resolution panchromatic imagery include an automatic segmentation approach based on watersheds [3]. Here the image is interpreted as a surface with the height proportional to the pixel intensity. The minima of this surface are the locations where water would settle if it were poured onto the surface. Watersheds typically form an over-segmentation of the image, so road candidate regions are then determined by an elongatedness calculation that detects rectangular regions. Finally, these candidates are merged with adjacent regions based on continuing the structural properties of the road. In [6], a semi-automated approach is presented, which involves tracing a road utilizing a Kalman filtering approach. Given an initial point and direction from an operator, the algorithm first extracts a cross-road profile. Then it takes progressive steps in the provided direction of the road, comparing the cross-road profile with the initial profile and adjusting the location and curvature estimates of the road, which allows for tracing of curved roads. The Kalman filter used in the update the position estimate enables the algorithm to trace through obstructions by ignoring profile information when it deviates too far from the initial profile by using the existing orientation and curvature estimates of the road in determining the road location.

Proceedings of the 31st Applied Imagery Pattern Recognition Workshop (AIPR'02)
0-7695-1863-X/02 $17.00 © 2002 IEEE

In this paper, we propose a semi-automated method for the extraction of roads from high resolution (1-meter) pan-sharpened multispectral IKONOS imagery. An operator provides an initial seed point on the road of interest, then the region is extracted using a level set method. Level set methods provide a powerful and efficient technique for performing feature extraction by casting the region extraction problem as a boundary evolution problem, where topological changes can be easily handled, and *a priori* domain knowledge (e.g., structural constraints) may be incorporated into the evolution process. The level set method tracks the evolution of a boundary front which is moving with a speed function that is normal to the boundary curve. This speed function can depend on local or global properties of the evolving boundary or on external forces. In our work, we use several image geometric and spectral properties of road networks to facilitate the extraction process within the level set framework. Further analysis through iterative smoothing refines the extracted region to accurately estimate the road centerline despite the presence of cars on the road, changes in the pavement or surface properties of the road, or obstruction resulting from foliage or shadows cast on the road by neighboring trees.

Our level set method for road extraction provides the following advantages:

- Exploits information from all spectral bands as well as the higher resolution panchromatic imagery.
- Only a single seed point is required by the human operator, no road directional information is needed.
- Capable of extracting a network of roads from a single seed point (i.e., not just a single road), thus requiring minimal operator assistance.
- Easily extracts curved roads, not just straight-line road segments.
- Able to extract around road obstructions such as cars, crosswalks, trees, etc...
- Time-efficient extraction (e.g., a 20K pixel region is extracted in only 484 milliseconds).

2. Multispectral Feature Extraction

Current IKONOS satellite sensors provide 4-meter resolution multispectral data in the blue, green, red, and near-infrared bands, and 1-meter resolution panchromatic data, which is the highest resolution available in commercial satellite imagery today. The 11-bit radiometry supports the development of robust feature extraction algorithms by providing discernable detail even in the presence of shadows, highlights, and low-contrast scenes. The multispectral bands approximate LANDSAT bands one through four, therefore algorithms developed for the spectral analysis of LANDSAT data are applicable to IKONOS imagery with little or no modification.

2.1. Pan-Sharpening Procedure

The colorizing of 1-meter panchromatic pixels with the 4-meter multispectral data produces a pan-sharpened image, which has been shown to be very effective in both human visual interpretation of satellite imagery, as well as automated approaches for the extraction of cartographic features from LANDSAT imagery. Hence, we have developed a pan-sharpening procedure for the efficient exploitation of information provided by all image bands.

Each pixel in the panchromatic image is mapped to a non-integral location in the multispectral image by multiplying by the resolution ratio, which in this case is $1/4$. The four-band source color is then calculated by bilinearly interpolating the pixel values in each band. The four-band color is then viewed as a vector in color space, which is first normalized, and then multiplied by the intensity given in the panchromatic image. This conversion can be summarized as

$$\hat{c}(x,y) = I(x,y)\frac{\bar{c}(x,y)}{\epsilon + |\bar{c}(x,y)|} \tag{1}$$

where $\bar{c}(x,y)$ is the interpolated color value and ϵ is a small positive constant, used to compensate for cases where $|\bar{c}(x,y)|$ is small. The pan-sharpened image provides the road extraction algorithm with a rich source of information, that preserves the details provided by the panchromatic data, while allowing for the discrimination of vegetation, water and manmade surfaces.

2.2. Multispectral Analysis

Our road extraction approach makes use of the multispectral information by generating pixel classification masks of non-road-like pixels, which are incorporated into the level set evolution function to hinder growth into vegetation or water regions. Since the IKONOS multispectral bands approximate LANDSAT bands one through four, we can adopt classical spectral classification techniques based on a set of band ratio measures to generate the masks. The first pixel classification mask is based on the Transformed Vegetative Index (TVI), which is computed as

$$TVI = \sqrt{\frac{NIR - RED}{NIR + RED} + \frac{1}{2}} \tag{2}$$

The values of NIR and RED are the pixel values from the near-infrared and red bands. This index is based on the observation that vegetation has a high reflectance in the near-infrared band, while the reflectance is lower in the red band. The TVI value for manmade surfaces is small, while it is large for vegetation regions. This index is less sensitive to atmospheric conditions than other indices, therefore it has been widely used for vegetation monitoring. Figures 1.(a)-(c), 2.(a)-(c), provide examples of the classification masks generated using the TVI measure.

The second classification mask is generated by computing the Water Index (WI), which is calculated as

$$WI = \frac{NIR}{GRN} \tag{3}$$

the ratio of the pixel values of the near-infrared (NIR) and green (GRN) image bands. Reflectance from water surfaces is relatively low, and for wavelengths beyond the visible red band it is nearly zero. Therefore, the WI value is small when the region is likely to be a water body, and it is large for other types of surfaces. Figure 1.(d)-(f), shows an example of the classification mask generated using the WI measure.

3. Road Extraction Using Level Set Methods

Level set methods have been used extensively in applications involving boundary evolution, because the framework allows for cusp, corners, and automatic topological changes. These characteristics make the level set framework a logical choice for road extraction, since the extracted road boundary may be disjoint due to the presence of cars, possess sharp corners due to intersections, and can arbitrarily change its topology (e.g., roads may merge at any point along the road). Therefore, we pose the road extraction problem as a boundary evolution problem within the level set framework.

3.1 Level Set Formulation

The classical level set boundary is defined as the zero level set of an implicit function Φ defined on the entire image domain. The dimensionality of the level set function Φ is one higher than the evolving boundary, which in this case for a 2-D image domain is a 3-D surface. The level set method tracks the evolution of a front that is moving normal to the boundary with a speed $F(x,y)$. The speed function can be dependent on the local or global properties of the evolving boundary or from forces external to the boundary. The level set function $\Phi : \Re^2 \rightarrow \Re$ is defined such that the location of the boundary, Γ and the region enclosed by the boundary, Ω, are functions of the zero-level set of Φ, namely

$$\Gamma(x,y) = \begin{cases} 1 & \Phi(x,y) = 0 \\ 0 & \text{else} \end{cases}$$
$$\Omega(x,y) = \begin{cases} 1 & \Phi(x,y) \leq 0 \\ 0 & \text{else} \end{cases} \tag{4}$$

Generally, $\Phi(x,y)$ represents the *signed distance function* to the front. That is,

$$\Phi(x,y) = \pm d(x,y) \tag{5}$$

where $d(x,y)$ is the smallest distance from the point (x,y) to the boundary, and the sign is chosen such that points inside the boundary have a negative sign and those outside have a positive sign.

The function Φ is initialized based on a signed distance measure to the initial front. In the case of a single point this is simply the euclidian distance to that point. The evolution of the boundary is defined via a partial differential equation on the zero level set of Φ

$$\frac{\partial \Phi}{\partial t} = -F|\nabla \Phi|. \tag{6}$$

Finally, level set methods may be easily adapted to a discrete grid implementation for efficient processing. See [4] for a more detailed description on level set methods and their applications.

3.2. Application to Road Extraction

The level set framework has been adapted to a number of applications beyond front propagation of physical phenomena, including the denoising of images, object recognition, and shortest path identification. Key to the success of these approaches is determining the appropriate stopping criteria for the evolution. The level set will continue to propagate as long as the speed function $F(x,y) > 0$. Therefore, the speed function must be properly designed such that $F(x,y) \rightarrow 0$ as the boundary approaches the desired location. For the road extraction problem, it is desirable for $F(x,y) \rightarrow 0$ at the edges of the true road boundary. To achieve this, we exploit the following characteristics of roads: (1) spectral uniformity; (2) consistent textured surface, which is often smooth; and (3) good contrast with surrounding environment. This leads to a speed function $F(x,y)$ constructed as follows:

$$F(x,y) = (we^{-\frac{1}{2}(\hat{c}(x,y)-\bar{\mu}_o)\sum^{-1}(\hat{c}(x,y)-\bar{\mu}_o)^T} + (1-w) \quad e^{-|H(x,y)-H_o|}) * \frac{1}{1+|\nabla I(x,y)|^p}. \tag{7}$$

The first term in the product (7) expands the boundary to include image points based on spectral and textural similarity. It is broken down into two components. One term measures the spectral similarity of an image point (x,y) to that of the seed point, which reflects the property that roads are comprised of pixels having fairly similar spectral intensities. Here, $\hat{c}(x,y)$ denotes the spectral values at image location (x,y), and $\bar{\mu}_o$ represents the vector of spectral intensities of the seed point, which is compared to all other points. When the image value is close to $\bar{\mu}_o$, the resulting intensity term is close to 1, whereas when the image value is far removed from $\bar{\mu}_o$, the intensity term tends toward 0.

The other term measures the textural similarity between an image point (x, y) and the seed point, thus modeling the consistent smoothness of the surface of roads by matching the local entropy values. The function $H(x, y)$ is the entropy of the image at point (x, y), and H_o is the entropy of the seed point. To form the measure of entropy $H(x, y)$, the intensities in the local neighborhood of (x, y) are histogrammed into $N = 64$ bins, then the histogram is smoothed using a Gaussian ($\sigma = 1.0$), and used to compute the standard entropy calculation

$$H(p) = \sum_{i=0}^{N-1} p_i log p_i \tag{8}$$

where p_i represents the value in the ith bin after convolution with a Gaussian. As with the intensity term, a close match in entropy will result in a value close to 1. As the dissimilarity increases, the term will decrease toward 0. The parameter w determines the relative weightings between the intensity and entropy terms.

The second term in the product (7) is the gradient-based term, which seeks to stop the evolution process at the true road boundaries. Often a road will have good contrast with its surrounding areas, thus resulting in strong edges between the road boundaries and other regions. The ∇I, or gradient of the image, is large when there is high variation between neighboring values, thus denoting the presence of an edge. Therefore, the resulting term is such that when the gradient is small the term evaluates near 1, but decays towards 0 as the gradient increases. The additional parameters p and the diagonal covariance matrix $\sum$, allow for tuning the sensitivity of the two terms, and are a function of the noise and contrast present in the image.

The classification masks generated by the multispectral analysis are incorporated into the speed function $F(x, y)$ by adding two additional terms designed to slow the growth into non-road-like regions. Thus, for road extraction from multispectral imagery, the speed function becomes

$$\hat{F}(x, y) = F(x, y) * e^{-|TVI(x,y)-0.8|} * e^{-|WI(x,y)-0.9|} \tag{9}$$

where $F(x, y)$ denotes the original speed function (7). The two additional factors tend toward 0 as the image moves into regions with large likelihood of being vegetation or water. Since the two indices are based on ratios of the various bands, predetermined thresholds can be effective across all multispectral data.

One additional characteristic of roads which is not captured directly by (9) is the smooth continuity or slow varying nature of roads. Portions of the road may be rejected as a result of variations due to illumination or obstacles such as cars, trees, or the lane markers. The level set framework provides a way of regularizing erroneous rejections due to obstacles such as these. Let $\kappa(x, y)$ be the curvature of the boundary. When the boundary is concave, κ will be negative, and when the boundary is convex, κ will be positive. Therefore, $F(x, y)$ should only be nonzero when κ is negative, leading to a concavity removal equation given by

$$F(x, y) = -min(\kappa(x, y), 0) \tag{10}$$

Note that (10) is largest for points of highest curvature. Therefore, those parts of the boundary with higher concavity will be removed first. If this evolution is continued until a fixed point solution is reached, all concave regions of the boundary will be removed. Some concavity is acceptable, such as the concavity resulting from a curved road. The "local" concavity resulting from small omitted regions will have a large concavity, causing those regions to be removed at first. Therefore, evolving the front under (10) for a finite number of cycles will smooth out local irregularities without changing the global properties of the boundary.

The benefit of the level set formulation is that the curvature κ can be computed directly from the level set function Φ via

$$\kappa = \frac{\Phi_{xx}\Phi_y^2 - 2\Phi_x\Phi_y\Phi_{xy} + \Phi_{yy}\Phi_x^2}{(\Phi_x^2 + \Phi_y^2)^{3/2}} \tag{11}$$

where Φ_x denotes the partial derivative of Φ with respect to x, Φ_{xy} denotes the partial derivative of Φ with respect to x and y, and so on.

The equation in (9) always reflects an outward growth and as such can be formulated in the level set framework of fast marching methods [4], leading to a very efficient implementation. The smoothing outlined in (10), however, cannot be implemented in such a manner, owing to the variation of the local curvature computation κ over time. Instead of resulting to a full level set implementation, it is possible to improve our execution time by evolving the level set only in the neighborhood of the boundary of the region. Our implementation is based on this two stage extraction process in order to provide fast and efficient extraction of roads.

3.3. Road Segment Analysis

After a region has been extracted and smoothed, the centerline is estimated by a skeletonization procedure, which preserves the connectivity of the original extracted domain. Junction analysis is then performed to ensure that the road centerlines meet at appropriate angles within intersections. This is accomplished by computing the angular texture signature at points lying within the intersection. For each angle, we compute the variance over a rectangular region extending from the pixel in the angular direction. At the true intersection point, the angular texture signature will exhibit a distinct set of peaks indicating the number of roads that meet at that point. Intersecting road segments having the same road properties are then automatically linked.

4. Experimental Results

We have tested our road extraction algorithm on several IKONOS image samples depicting the Atlanta and Baltimore areas. The samples contained roads in both urban and rural settings as shown in Figure 3.(a)-(d), and were selected to demonstrate the algorithm's capabilities. Case A is an extraction of a curved road having a weak boundary. Case B is a busy highway and demonstrates extraction around objects. Case C is an extraction of a complex road network with a single seed point provided. Case D is a three-way intersection obstructed by a crosswalk. Case B is initialized using two operator-provided seed points due to the fact that the two highways are not connected and do not form a network. All other cases are initialized using a single seed point. These seeds are indicated by the plus signs in the images shown in Figure 3.(a)-(d).

All test cases were executed on a Pentium 4, 1.6 GHz system with 512 MB of RAM. The results are shown in Figure 4 with the road boundary and centerline delineated. We compared our road estimates with boundaries that were manually hand-annotated or "ground truthed" (GT). For quantitative comparison, we compute the following accuracy and precision measures

$$Accuracy = \frac{|GT \cap Estimate|}{|GT|} x100\% \qquad (12)$$

$$Precision = \frac{|GT \cap Estimate|}{|Estimate|} x100\%. \qquad (13)$$

In both of these equations, $|\cdot|$ represents the size, in pixels, of the given domain. Accuracy will be large when estimate includes all parts of the ground truth and precision will be large when there are no spurrious regions included in the estimate that are not in the ground truth. In the case of roads, a centerline estimate is also computed by the algorithm, and the RMS distance of the estimate to the hand-annotated centerline is given. The numerical results presented in Table 1 show good performance of the algorithm with respect to both region extraction and centerline estimation.

Test Image	Size Pixels	Acc. %	Prec. %	Center RMS	Time (ms)
A: Curve	2,377	97.0	96.8	0.55	67
B: Highway	20,135	99.2	95.2	0.30	484
C: Network	5,289	99.4	95.8	0.58	120
D: Crosswalk	3,611	97.6	96.3	0.41	219

Table 1. Road Extraction Performance

5. Conclusion

In this paper, we presented a semi-automated road extraction algorithm based on level set methods. The operator initializes the algorithm by providing a single seed point, then the level set based region growing algorithm extracts the road network or segment. Our level set algorithm for road extraction provides several benefits:(1) the ability to extract roads of varying topology; (2) requires minimal human intervention; (3) exploits the information available from all imagery; (4) capable of dealing effectively with cars, crosswalks and other road obstructions; and (5) the algorithm is computationally efficient. We are investigating methods for enhancing the robustness of the extraction process, further multispectral analysis for characterizing regions, and the automatic discovery of good road seed points to further minimize the amount of human intervention.

References

[1] D. German and B. Jedynak, "An Active Testing Model for Tracking Roads in Satellite Images", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 18, no. 1, Jan. 1996, pp. 1-14.

[2] M. Fischler, et al, "Detection of Road and Linear Structures in Low Resoltuion Aerial Imagery Using a Multisource Knowledge Integration Technique", *Computer Graphics and Image Processing*, vol. 15, 1981, pp. 201-223.

[3] H. Y. Lee, et al, "Towards Knowledge-Based Extraction of Roads from 1m-resolution Satellite Images", *Proc. of the IEEE Southwest Symposium on Image Analysis and Interpretation*, Austin, Texas, April 2000, pp. 171-176.

[4] J. A. Sethian, "Level Set Methods and Fast Marching Methods", Cambridge University Press, United Kingdom, 1999.

[5] C. Steger, "An Unbiased Detector of Curvilinear Structures", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 2, Feb. 1998, pp. 113-125.

[6] G. Vosselman, and J. Knecht, "Road Tracing by Profile Matching and Kalman Filtering", *Proc. of the Workshop on Automatic Extraction of Man-Made Objects from Aerial and Space Images*, Birkhaeuser, 1995, pp. 265-274.

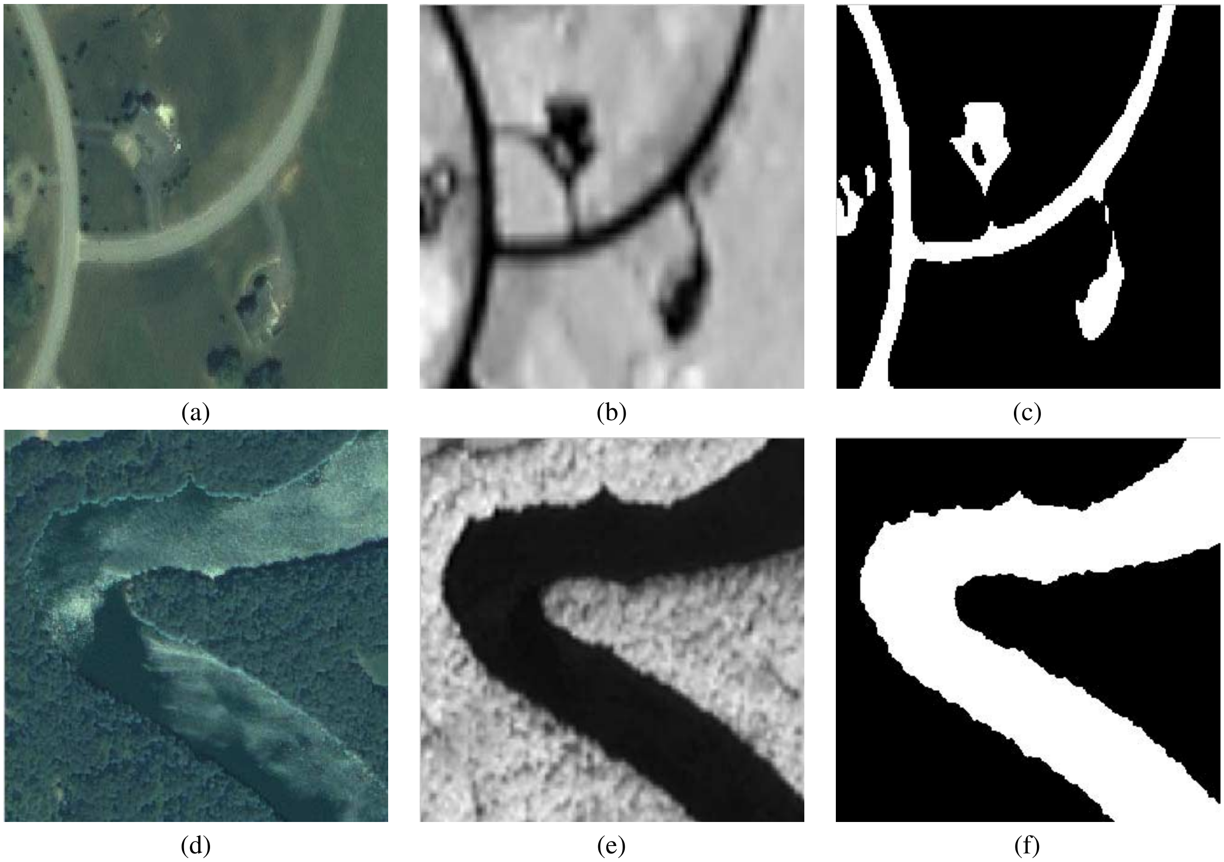

Figure 1. Multispectral Analysis: (a) Pan-Sharpened Road Image; (b) TVI Index Map; (c) Spectral Classification Mask (White - Manmade Regions); (d) Pan-Sharpened Water Image; (e) WI Index Map; (f) Spectral Classification Mask (White - Water Body Regions)

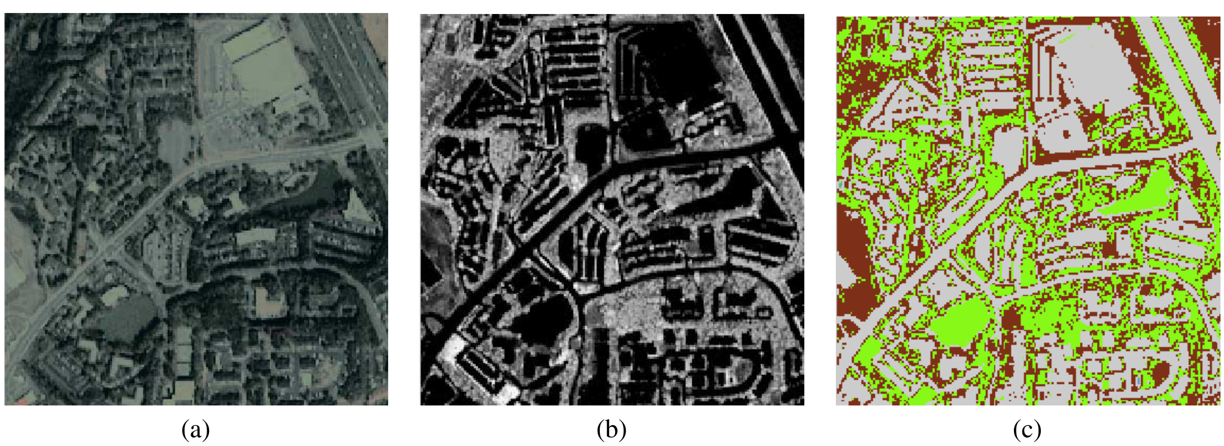

Figure 2. Multispectral Analysis: (a) Pan-Sharpened Urban Image; (b) TVI Index Map; (c) Spectral Classification Mask (Green - Vegetation; Brown - Soil; Gray - Manmade)

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

Figure 3. Test Cases: (a) Curved Road (b) Busy Highway (c) Road Network (d) Crosswalk

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

Figure 4. Extraction Results