ROAD TRACKING BY MAXIMIZATION OF MUTUAL INFORMATION

Xiaoying Jin
Scott Paswaters
Research Systems Inc.
4990 Pearl East Circle
Boulder, CO 80301
jinxy65201@yahoo.com
scottp@rsinc.com

Curt H. Davis
Center for Geospatial Intelligence
Department of Electrical and Computer Engineering
University of Missouri-Columbia
Columbia, MO 65211
DavisCH@missouri.edu

ABSTRACT

Automated and semi-automated geospatial feature extraction has attracted great interest from academic and industry researchers in the last decade. In this paper, we present a novel semi-automated algorithm for tracking road centerlines from medium and high-resolution remote sensing imagery. Roads are modeled as continuous and curvilinear features with similar intensity profiles perpendicular to the road centerlines. Our algorithm starts from a user-defined road seed point located on the road's center and tracks the road centerline by searching similar image subsets along the road direction. A square window around the seed point is used as a road template. The initial position for the next road point is estimated by linear extrapolation. It is then tuned toward the road centerline by maximizing the mutual information between the template and the image subset around the tracked point. Mutual information was first defined in information theory and has been used in inter and intra-modality image matching and registration since the 1990s. Previous research on semi-automated road extraction used least squares template matching (LSTM) techniques. This paper demonstrates that mutual information is also a useful similarity metric for road tracking. Experiments performed on a variety of test sites demonstrate that the proposed algorithm is more robust than algorithms using other similarity metrics. Specifically, the algorithm proposed here is less sensitive to variations in scene illumination and surface material changes.

INTRODUCTION

Automated and semi-automated feature extraction from remotely sensed imagery has attracted great interest from academic and industry researchers in the last decade. Accurate and up-to-date road information is essential for geographic information system (GIS) applications as well as transportation and urban planning. Research has been stimulated by the increased availability of high-resolution remote sensing imagery in recent years from a variety of airborne and satellite sensors. With digital remote sensing imagery at 1-m or less image resolution, roads can be readily seen from the imagery. Since manual extraction of road networks from imagery is very time consuming, automated and semi-automated methods have the potential to improve the speed and utility for this important application.

Road extraction strategies are usually classified into two categories according to the degree of human interaction: semi-automated and fully-automated extraction. There is a trade-off between user interaction and automation. Usually fully automated algorithms produce incomplete results with some false extractions so that manual post processing is inevitable (Quackenbush, 2004). Due to the complexity of the automation problem, human operators still play an important role in extracting roads from imagery. Semi-automated methods were considered a good compromise, combining the speed and accuracy of a computer programs with the interpretation skill of a human operator. For semi-automated road extraction, initial seed points, sometimes with directions, must be input to an algorithm that attempts to connect the seed points using various search path criteria. This can be built into an integrated road extraction system if an algorithm can reliably find road seeds from the imagery (Jin and Davis, 2005). In next section, a short review of previous work on semi-automated road extraction is given.

PREVIOUS WORK

A significant amount of semi-automated road extraction research has been done. Since fully automated methods for road mapping were far out of reach for early research, semi-automated methods were considered a good compromise, combining the speed and accuracy of a computer programs with the interpretation skill of a human operator.

Semi-automated methods can be roughly classified into two categories. One category is called road tracking or following. A human operator must provide a starting point and direction for a trackable feature such as a road. Then, the algorithms may track or follow the road from the starting point based on some radiometric, geometric, or topological characteristic to trace the remaining road sections. Another category is called road linking. Here, a human operator identifies the object of interest (roads in this case) from a digital image and selects a few seed points that are coarsely distributed. Then, linear features are extracted precisely and automatically by computer algorithms to link the points based on radiometric, geometric, and/or topological measures.

Template or profile matching is an algorithm type used in road tracking strategies. Usually, a starting point and direction are input by a human operator. The method makes use of some initial information about the road, i.e., the position and direction of a starting point, and approximate road width. In Quam's road tracker (Quam, 1978), roads are traced in the high-resolution imagery with a ground sample distance (GSD) of 1-3 m. A parabolic extrapolation path model was used to predict the next point based on a list of recent road points. A set of intensity profiles near the predicted position are matched with the reference surface model. A surface model is an array of intensity values sampled from the image in the direction perpendicular to the road. The best matching position is chosen as the actual road position. If the cross-correlation peak is poor and the match failed for several successive steps, the road tracker stops. McKeown and Denlinger proposed a cooperative method for road tracking in aerial imagery. A surface model road tracker and an edge tracker work independently based on different cues. An intermediate level process is used to evaluate the success of each tracker. If one tracker fails, it can be started from the model generated by other successful trackers (McKeown and Denlinger, 1988). Vosselman and Knecht used profile matching and Kalman filtering for road tracking. In their method, road positions are computed by matching the average gray value profile of a reference road segment with profiles taken from the image. The road parameters are estimated by a recursive Kalman filter. By utilizing the predictive power of the Kalman filter the road tracker is able to continue following the road despite temporary failures of the profile matching (Vosselman and Knecht, 1995).

Previous research on road tracking often employed least squares template matching (LSTM) techniques. LSTM has rigorous theory and methods for quality control and result assessments (Gruen, 1985), and it has been successfully used in semi-automated linear feature extraction (Hu, 2004; Kim, 2004). Hu *et al.* used a piecewise parabolic model as their road trajectory model. Using a dual step-edge finder, deformable road templates self-adapted to slight changes of road width and variations in dual edge intensities (Hu, 2004). Kim *et al.* traced the road a small step ahead at each iteration and used a piece-wise straight line as the road trajectory model. Both of these previous techniques used LSTM. The initial road direction was estimated by the orientation of the line connecting two consecutive input seed points. Let $f(x,y)$ be the gray values of the template that represents the gray pattern of the road and $g(x,y)$ be the observation gray values of the kernel in the current estimated feature position. Correlation is ideally established if

$$f(x,y) = g(x,y)\,. \tag{1}$$

The objective function to be minimized in LSTM is the sum of the squared differences between intensity values of the template and the observation.

In summary, a basic assumption of road tracking methods is that the radiometric properties and the direction of the road will not change abruptly, and this is satisfied by most local sections of the road network. The algorithms start from a user-defined road seed point located on the road center and track the road centerline by searching similar image subsets along the road direction. The similarity metric optimized in previous research includes cross-correlation and least squares correlation matching scores. In the next section, we propose a new similarity metric - mutual information - for road tracking.

METHOD

In this section, we present a novel semi-automated algorithm for tracking road centerlines from medium and high-resolution remote sensing imagery. Roads are modeled as continuous and curvilinear features with similar intensity profiles perpendicular to the road centerlines. Our algorithm exploits a road tracking approach. It starts from a user-defined road seed point located on the road center and tracks the road centerline by searching similar image subsets along the road direction. A square window around the seed point is used as a road template. The initial position for the next road point is estimated by linear extrapolation. It is then tuned toward the road centerline by maximizing the mutual information between the template and the image subset around the tracked point.

Road Template and Trajectory Model

Humans discriminate roads from imagery based on their spectral and spatial characteristics. In medium and high resolution images, roads are often modeled as continuous and elongated homogeneous regions with nearly constant width. Hence, they have similar intensity profiles perpendicular to the road centerline. In addition, road curvature has an upper bound so a road centerline can be approximated by a straight line over a short distance. Here we use a piecewise straight line as our road trajectory model. Given the position and direction of starting road seed point located on a road centerline, the initial position for the next road point over a short distance d is estimated by linear extrapolation. We use a square window around the seed point as a road template. The algorithm tunes the position of next road point toward the road centerline by maximizing the similarity metric between the template window and the target window around the tracked point.

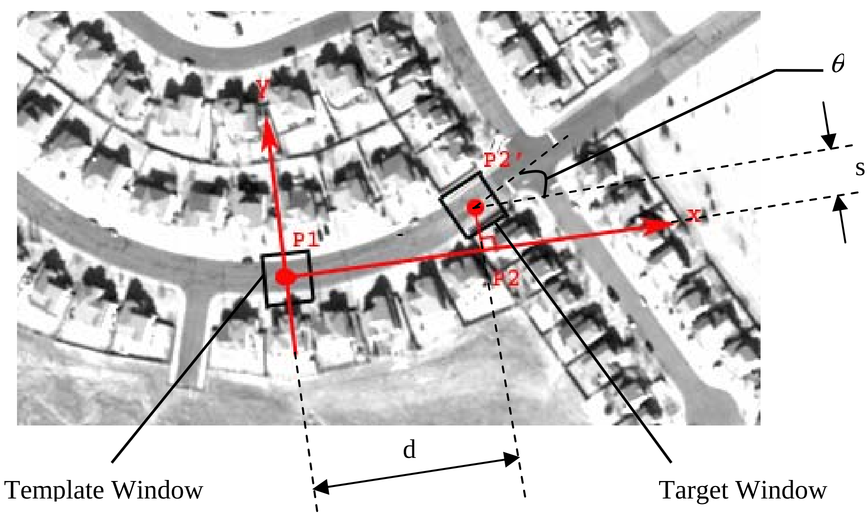


Figure 1. Road template and trajectory model.

The road template and trajectory model is shown in Figure 1. We define the origin of the tracking coordinate system as seed point $P1$ and the x axis pointing along the road orientation. The next point $P2$ is estimated by linear extrapolation one step d ahead. $P2$ is then tuned toward the real road center point $P2'$ by maximizing the similarity metric between the template window and target window. In this approach, there are two free parameters s and θ that can be tuned by the algorithm. We limit the update direction of target position to be along the direction normal to the road orientation, where s is the offset along the update direction and θ is the counter-clockwise rotation angle between the target window and road orientation. We can model the similarity transformation between the template and target windows in the tracking coordinate system as:

$$x_{target} = x_{template} \cos\theta - y_{template} \sin\theta + d \tag{2}$$

$$y_{target} = x_{template} \sin\theta + y_{template} \cos\theta - s \tag{3}$$

where $(x_{template}, y_{template})$ is the point coordinate in the template window and (x_{target}, y_{target}) is the coordinate of corresponding point in the target window. After the estimated point is tuned toward the road centerline, the road template and tracking coordinate system is updated using the new road centerline point $P2'$. The extraction continues until the similarity metric is below a certain value or it approaches the next seed point.

Mutual Information Similarity Metric

Mutual information was first defined in information theory. It measures the statistical dependency between two random variables or the amount of information that one variable contains from the other. Mutual information has been used in inter and intra-modality image matching and registration since the 1990s and represent the leading technique in multi-modal image registration (Cole-Rhodes, 2003; Viola, 1997; Zitova and Flusser, 2003). Mutual information between two random variables X and Y is given by

$$\begin{aligned} I(X,Y) &= H(X) + H(Y) - H(X,Y) \\ &= H(X) - H(X/Y) = H(Y) - H(Y/X) \end{aligned} \tag{4}$$

This definition is based upon entropy (*H*): $H(X)$ and $H(Y)$ is the entropy of variable X and Y, respectively and $H(X,Y)$ is their joint entropy. $H(X/Y)$ is the conditional entropy of X given Y (Papoulis, 1984). The entropy $H(X)$ is a measure of the amount of uncertainty for the random variable X. $H(X/Y)$ is the amount of uncertainty left in X when Y is known. Hence, mutual information $I(X,Y)$ is the amount by which the uncertainty about X decreases when Y is given, or, equivalently, the amount of information that X contains from Y. When X and Y are independent, the mutual information is zero. Mutual information is maximized when two random variables are highly dependent or even equivalent.

In image registration applications, pervious research demonstrated that mutual information enables one to extract an optimal match with a much better precision than cross-correlation (Cole-Rhodes, 2003). Mutual information has not been used previously in road extraction research up until now. Here we propose to use mutual information as a similarity metric during road tracking. The similarity criterion here states that the mutual information of the image intensity values in template and target windows is maximized when similar road image subsets are found. Because no assumptions are made regarding the nature of the relationship between the image intensities in both windows, this criterion is very general and powerful.

If we represent X and Y as the image subsets at the template and target window, respectively, the probability distribution of intensity values in X and Y can be computed using the histograms of the two images, $h(x)$ and $h(y)$ respectively. Mutual information is then computed as

$$I(X,Y) = \frac{1}{N} \sum_{x} \sum_{y} h(x,y) \bullet \log\left(\frac{N \bullet h(x,y)}{h(x) \bullet h(y)}\right) \tag{5}$$

where N is the image window size, and $h(x,y)$ is the joint histogram of two image windows.

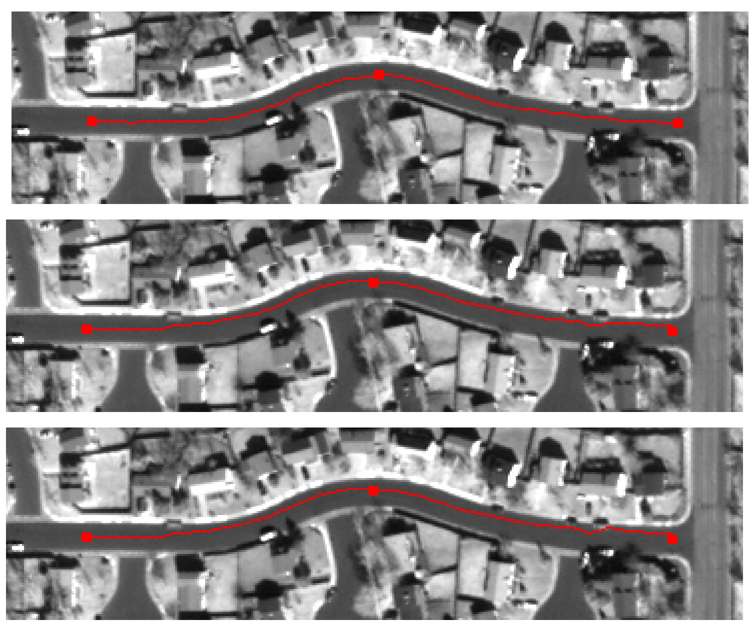
Search Strategy

Initially, we estimate the next point by linear extrapolation. The initial direction of the target window is along the road orientation. The algorithm then tunes the target window toward the road centerline by maximizing mutual information between the template and the target windows. There are two parameters to be tuned: s and θ, which represent the center location and rotation angle of the target window, respectively. Maximization of mutual information here is a 2-dimensional optimization problem. There are two categories of optimization algorithms: gradient-based and gradient-free strategies. The probability distributions required in the mutual information computation are estimated using the image histograms in Equation (5). The dependence of the mutual information measure on this discrete histogram makes the computation of its derivative complex. Hence, we chose a gradient-free optimization strategy. Powell's multi-dimensional direction set method is used to maximize $I(X,Y)$, using Brent's one-dimensional optimization algorithm for line minimizations (Press, 1992).

RESULTS AND DISCUSSIONS

The proposed road tracking algorithm was tested using a QuickBird high-resolution satellite image of Boulder, Colorado provided by DigitalGlobe. The panchromatic QuickBird data has a 0.7-m image resolution and is very suitable for this application. If we choose two close seed points located on the road centerline, the initial road orientation can be approximated by the orientation of the line connecting the two points. The extrapolation distance d was set to be 5 pixels. The width of the template and target windows was set to be one and a half times the road width. The algorithm starts from the first point and continues road tracking until it reaches the second point.

The idea of road tracking is to find similar road image subsets along the road direction. Cross-correlation and sum of squared differences (SSD) have been used in road tracking research before. In our algorithm, we use mutual information as a similarity metric. For comparison, we also used cross-correlation and SSD as two alternative similarity metrics between the template and target windows. Powell's multi-dimensional direction set method was used to maximize cross-correlation and minimize SSD, respectively. We compared our road tracking results using all three similarity metrics.



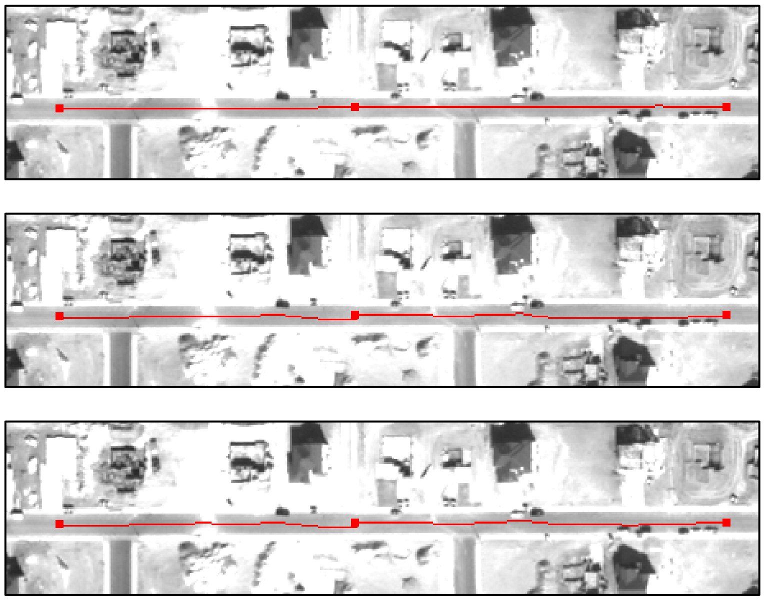
(a)

(b)

(c)

Figure 2. Road tracking results on QuickBird test site 1 using different similarity metrics: (a) mutual information, (b) SSD, and (c) cross-correlation.

We selected three test sites from the QuickBird image over Boulder, Colorado representing different conditions of road surface reflectance (Figs. 2, 3, and 4). In Figure 2, the road surface itself is smooth with small disturbances from surrounding trees, shadows, and parked vehicles. In Figure 3, the illumination condition changes in parts of the road surface. The road pavement changes around the road intersection in Figure 4. The road tracking results using different similarity metrics are shown in (a), (b), and (c) of the various figures. The seed points input by a human operator are denoted by red squared nodes. The tracked road centerlines are highlighted as red lines.

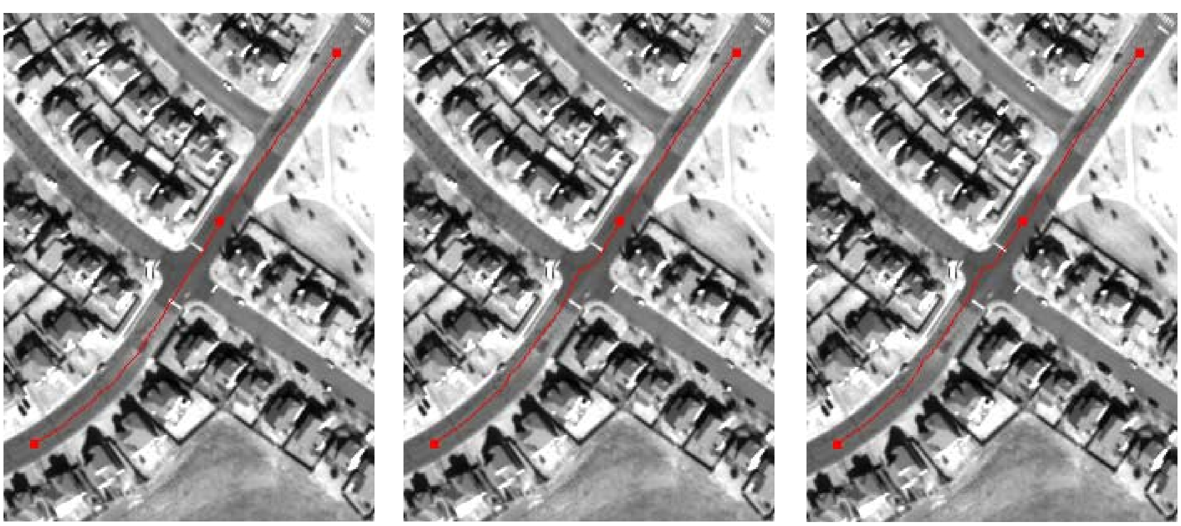


(a)

(b)

(c)

Figure 3. Road tracking results on QuickBird test site 2 using different similarity metrics: (a) mutual information, (b) SSD, and (c) cross-correlation.



(a) (b) (c)

Figure 4. Road tracking results on QuickBird test site 3 using different similarity metrics: (a) mutual information, (b) SSD, and (c) cross-correlation.

From the results, we see that road tracking using mutual information as a similarity metric is more reliable than the cross-correlation and SSD metrics in all three test cases. The tracking results using mutual information are robust to small disturbances along the road surface, variations of scene illumination and surface material changes. Road tracking using cross-correlation and SSD sometimes does not accurately track the road centerlines (Figure 2). The SSD similarity metric assumes that the intensity values of the template and target windows are similar. The cross-correlation metric relies on the assumption that the intensity values of two image windows are linearly correlated. Because no assumptions are made regarding the nature of the relationship between the image intensities in both windows and no constraints are imposed on the image content involved, mutual information is a very general and powerful similarity criterion and was successfully used in our road tracking strategy. The limited experiments here indicate a great potential for the developed method to be integrated seamlessly into practical linear feature GIS data collection process.

CONCLUSIONS

In this paper, we presented a novel semi-automated algorithm for tracking road centerlines from medium and high resolution remote sensing imagery. Our algorithm starts from a user-defined road seed point located on the road center and tracks the road centerline by searching similar image subsets along the road direction. This paper demonstrates that mutual information is a very useful similarity metric for road tracking. Experiments on a variety of test sites demonstrated that the proposed algorithm is more robust than algorithms using other similarity metrics with respect to variations of scene illumination and surface material changes. The current road tracking algorithm was applied to a single-band panchromatic input image. Future work would be extended the algorithm to work with multi-spectral data.

REFERENCES

Cole-Rhodes, A. A., K. L. Johnson, J. LeMoigne, and I. Zavorin (2003). Multiresolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient. *IEEE Transactions on Image Processing, 12*(12):1495-1511.

Gruen, A. (1985). Adaptive least squares correlation: A powerful image matching technique. *South African Journal of Photogrammetry, Remote Sensing and Cartography, 14*(3):175-187.

Hu, X., Z. Zhang, and C. V. Tao (2004), A robust method for semi-automatic extraction of road centerlines using a piecewise parabolic model and least square template matching. *Photogrammetric Engineering and Remote Sensing, 70*(12):1393-1398.

Jin, X., and C. H. Davis (2005). An integrated system for automatic road mapping from high-resolution multispectral satellite imagery by information fusion. *International Journal on Information Fusion - Special Issue: Fusion of Urban Remotely Sensed Features, 6*(4):257-273.

Kim, T., S. R. Park, M. G. Kim, S. Jeong, and K. O. Kim (2004). Tracking road centerlines from high resolution remote sensing images by least squares correlation matching. *Photogrammetric Engineering and Remote Sensing, 70*(12):1417-1422.

McKeown, D. M., and J. L. Denlinger (1988). Cooperative methods for road tracking in aerial imagery. In: *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, CVPR'88, pp. 662-672.

Papoulis, A. (1984). *Probability, Random Variables, and Stochastic Processes, second edition*. New York: McGraw-Hill.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling (1992). *Numerical Recipe in C: The Art of Scientific Computing, second edition*. Cambridge Univ. Press.

Quackenbush, L. J. (2004). A review of techniques for extracting linear features from imagery, *Photogrammetric Engineering and Remote Sensing, 70*(12):1383-1392.

Quam, L. (1978). Road tracking and anomaly detection in aerial imagery. In: *Proc. of Image Understanding Workshop*, pp. 51-55.

Viola, P. (1997). Alignment by maximization of mutual information. *International Journal of Computer Vision, 24*(2): 137-154.

Vosselman, G., and J. Knecht (1995). Road tracking by profile matching and Kalman filtering. In: *Automatic Extraction of Man-Made Objects from Aerial and Space Image*, Basel-Switzerland, pp. 265-274.

Zitova, B., and J. Flusser (2003). Image registration methods: a survey. *Image and Vision Computing*, *21*(11):977-1000.