I. Project Title

Automatic Road Extraction from Remotely Sensed Imagery

II. Lead Investigators

Timothy Haithcoat
Program Director
Geographic Resources Center
University of Missouri-Columbia
Room 104 Stewart Hall
Phone: (573) 882-2324
E-mail: haithcoatt@missouri.edu

III. Research Goals/Objectives/Rationale

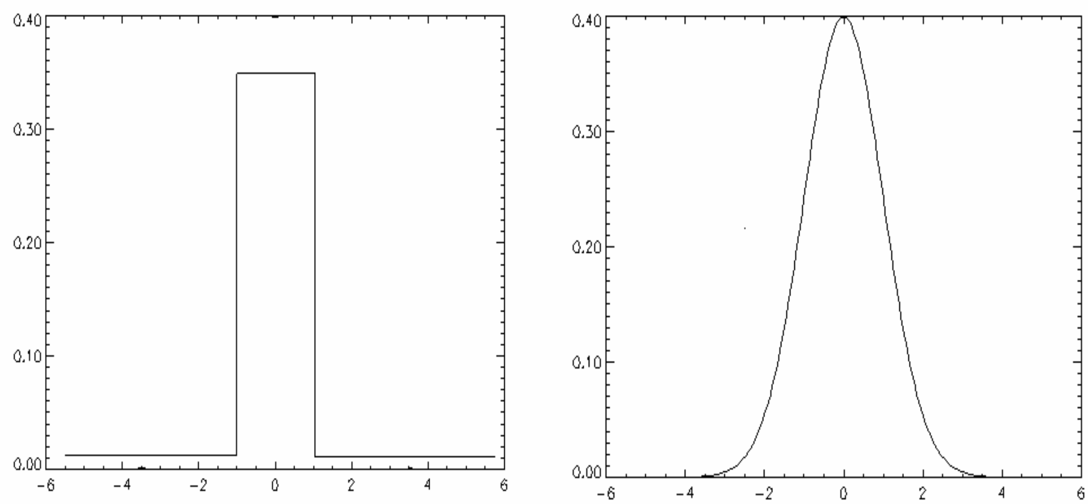
Today satellite and airborne remote sensing systems can provide large volumes of data that are invaluable in monitoring Earth resources and the effects of human activities. However, from a mapping perspective, research in remote sensing has been focused on land use/land cover classification. Fewer individuals are involved in cultural feature extraction research (the detection, identification, classification, and delineation of man-made features). Recently, remarkable progress has been achieved in digital (softcopy) photogrammetry. Now, digital elevation models (DEM) and digital orthophotographs can be automatically generated. However, the tasks of feature identification and cartographic delineation are still done manually. These tasks are time consuming, very labor intensive and therefore costly.

Innumerable public agencies and private companies require current and complete digital cartographic information. Typical map products cannot maintain currency because of the rapid pace of development. Remote sensing systems can provide current raw data in an image format to update these databases. What is needed is the development of an automatic, fast and reliable approach to extract cartographic information from this imagery. Feature extraction from remotely sensed imagery has special importance to map creation for developing countries as well as rural areas in developed countries. In these areas, there are often no current maps or for some, even no maps at all.

This research proposed an integrated approach for automatic road extraction from remotely sensed imagery by combining digital image processing, remote sensing and geographic information system (GIS) technologies. Roads are modeled as continuous single lines with bar-shaped or parabolic-shaped profiles in the direction perpendicular to the road. The roads are extracted based on radiometric and geometric properties. Then, further GIS operations are applied to obtain vector roads with higher cartographic quality. Landsat7 ETM+ data with 30 meters resolution, 1 meter USGS DOQ (Digital Orthophoto Quadrangle), 1 meter IKONOS imagery and 0.25 meter scanned aerial photography were used to test this approach. The results were evaluated through comparison to manually acquired road data. Several quality measures (Completeness, Correctness, Quality etc.) were used within the accuracy assessment. The extraction algorithms have been successfully ported to a desktop GIS environment. This integration of GIS and remote sensing technologies provides a promising approach for automatic GIS data collection, update and maintenance.

IV. Summary of Research Activities

This research models roads as linear features existing within grayscale imagery. The image is regarded as a function with roads having bar-shaped or parabolically shaped grayscale profiles in the direction perpendicular to the road (Figure 1).



(a) (b)

Figure 1. Road Profiles: (a) Bar-shaped (b) Parabolically shaped

The automatic road extraction process includes three basic procedures: image preprocessing, road extracting and GIS processing (Figure 2). Image preprocessing is conducted to facilitate or enable further operations. This process varies depending on the Image characteristics. For Landsat7 ETM+ data, a principal component analysis was performed to enhance the image content following selection of the PCA component exhibiting the best road contrast. For DOQ, IKONOS and scanned aerial photograph, the original images were re-sampled to a coarser resolution so that the roads were only a few pixels wide. The road extraction algorithms were then applied on these reduced resolution images. Images were convolved with derivatives of Gaussian kernels. A Hessian matrix was constructed and its eigenvalues and eigenvectors were calculated. After thresholding the second derivative image, a raster-to-vector conversion was conducted to extract the vector roads. GIS operations were then applied to refine the results both topologically and aesthetically.

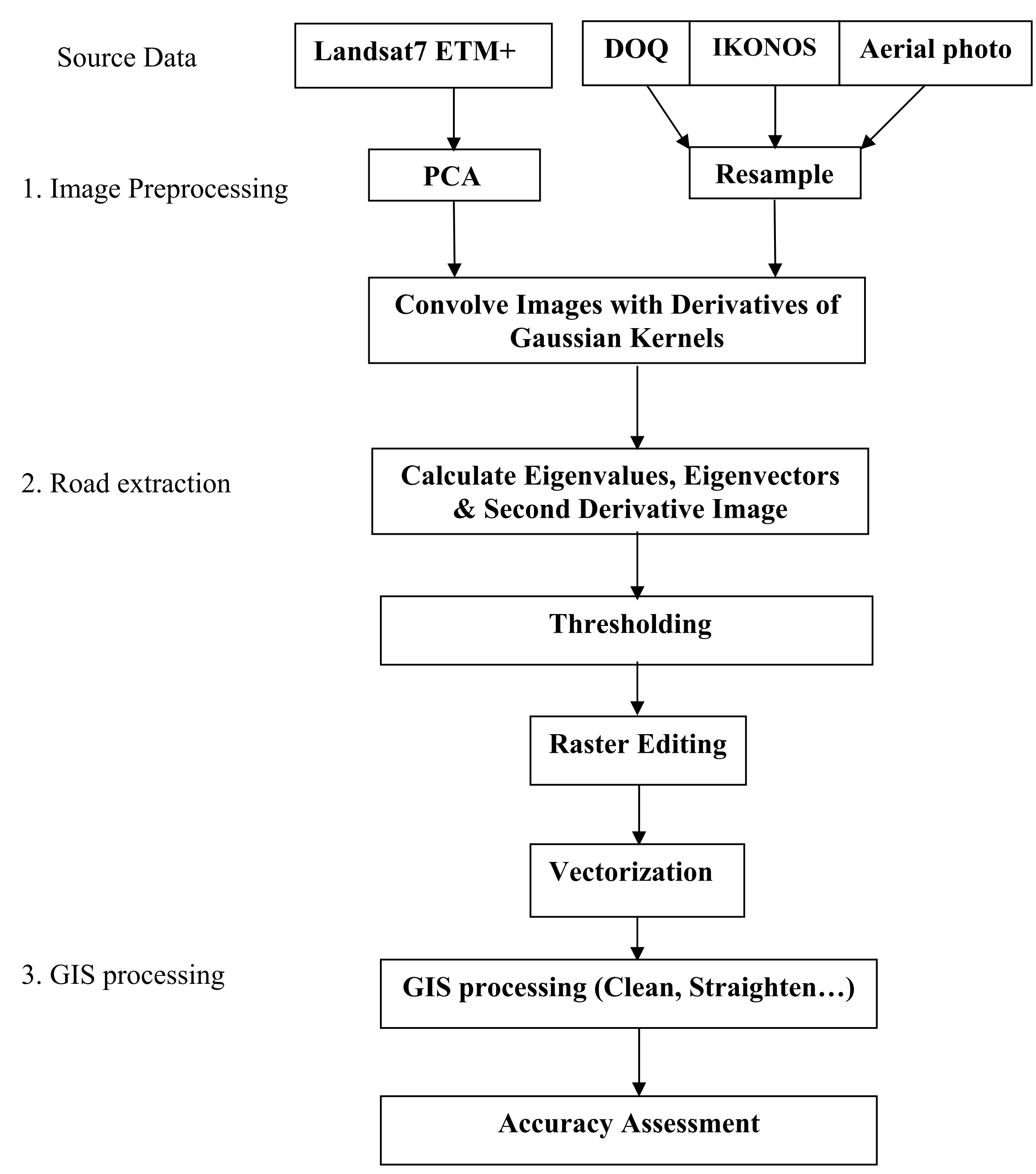


Figure 2. Automatic Road Extraction Scheme

Factors such as resolution and context play important roles in determining the success of image-based road extraction. Therefore, a matrix of resolution sub-images (30 meters Landsat7 ETM+, 1 meter IKONOS, 1 meter DOQ, and 0.25 meter aerial photograph) with forested, rural, suburban, residential, and urban areas was used to assess the robustness and adaptability of the algorithm.

The automatically extracted roads were compared with manually interpreted reference roads for this accuracy assessment. Because roads are linear features, it was possible to use all the reference data to conduct the accuracy assessment in a GIS environment. The following measures were used for this accuracy assessment:

- $$\textbf{Completeness} = \frac{\textbf{Length of matched extraction}}{\textbf{Length of reference}} \quad (1)$$
- $$\textbf{Correctness} = \frac{\textbf{Length of matched extraction}}{\textbf{Length of extraction}} \quad (2)$$
- $$\textbf{Quality} = \frac{\textbf{Length of matched extraction}}{\textbf{Length of extraction} + \textbf{Length of unmatched reference}} \quad (3)$$
- $$\textbf{RMSE} = \sqrt{\sum[(x_{extrcted} - x_{reference})^2 + (y_{extracted} - y_{reference})^2]/n} \quad (4)$$
- **Horizontal Accuracy = 1.7308 * RMS (Root Mean Square)** (5)

To calculate these measures, buffer zones were generated around the extracted roads as well as the reference roads. The chosen buffer width was set to approximately half of the actual road width on the image. Matched extraction roads were derived by intersecting the extracted roads with the buffered reference. Matched reference roads were derived by intersecting the reference with the buffered extracted roads. The RMS error and horizontal accuracy measure was obtained by calculating the actual distance between extracted road points and reference roads.

Figure 3 shows the results of road extraction from Landsat ETM+ image. Figure 4 shows an example of road extraction from a suburban DOQ image and Figure 5 shows the results from IKONOS image of a residential area.

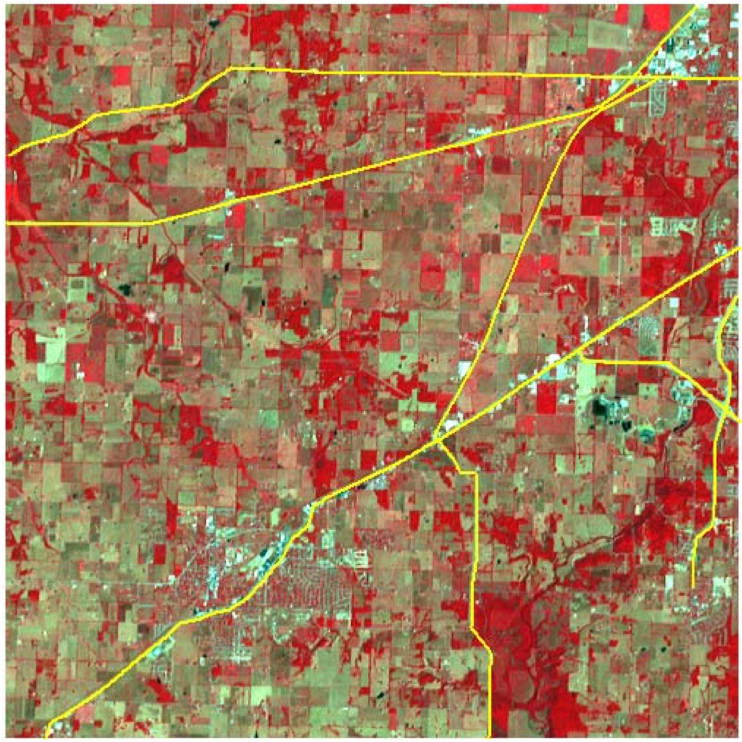


Figure 3. Extracted roads (yellow) from ETM+ image

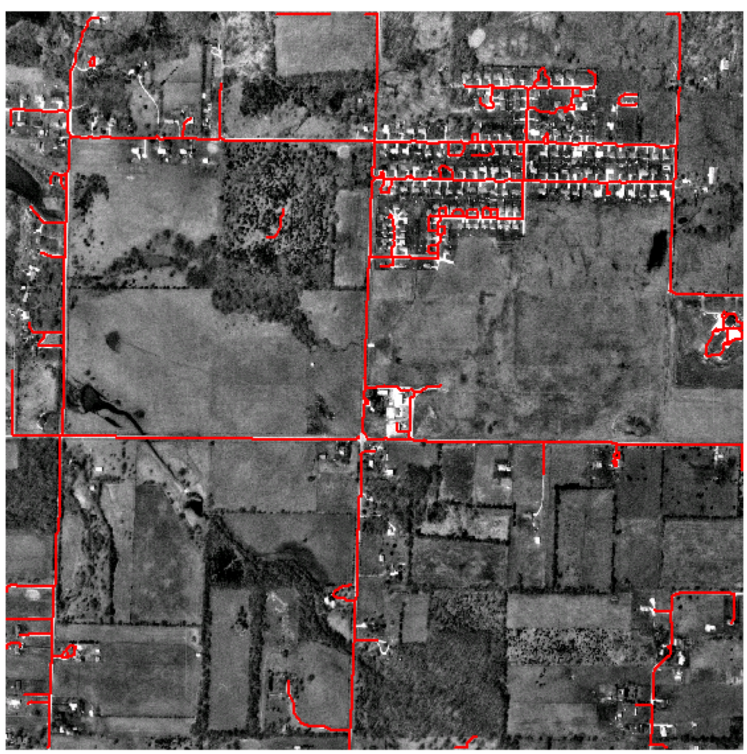


Figure 4. Extracted roads (red) from a DOQ image

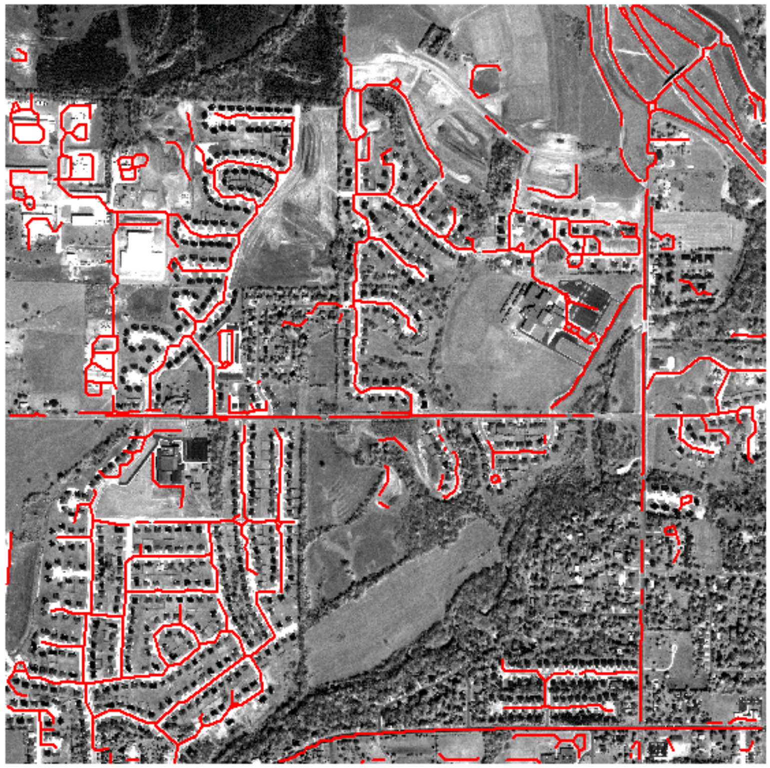


Figure 5. Extracted roads (red) from IKONOS image

The accuracy assessment for all tested images are listed in Table 1. Both Landsat7 ETM+ images were simple rural scenes and only major roads were tested for extraction. The algorithm generates good completeness (84% & 86%), correctness (95% & 91%) and quality (80% & 78%) statistics. Values of 89% and 95% completeness were achieved from rural and suburban DOQ images. The results were similar for the IKONOS imagery and aerial photography. We calculated the RMS error and horizontal accuracy measures only for the matched road extraction. All RMSE and horizontal accuracy measures were good for the matched roads (Table 1). Most values fell within the actual road width as represented as the buffer distance.

Images	Completeness (%)	Correctness (%)	Quality (%)	RMSE (m)	Horizontal Accuracy (m)
ETM+ Rural	84	95	80	13.37	23.14
ETM+ Rural	86	91	78	23.30	40.33
DOQ Rural	89	77	71	1.80	3.12
DOQ Suburban	95	66	63	2.00	3.46
DOQ Residential	73	41	35	2.10	3.63
DOQ Urban	60	37	29	3.70	6.40
IKONOS Rural	70	54	44	2.41	4.17
IKONOS Urban	83	66	58	3.08	5.33
Aerial Photograph	96	38	36	2.43	4.20

Table 1. Accuracy Assessment summary for 9 test images

There are several factors influencing the quality of the road extraction. These include spatial resolution, spectral information (contrast), and context (rural, urban). This research attempted to quantify these factors and examine their inter-relationships. The resolution, road density, number of intersections and contrast data for all images was calculated (Table 2).

Images	Resolution (Meters)	Road Density	Number of Intersections	Contrast
ETM+ Rural	30	0.146	0.036	135
ETM+ Rural	30	0.310	0.146	111
DOQ Rural	1	4.57	48	134
DOQ Suburban	1	4.42	81	138
DOQ Residential	1	11.39	147	71
DOQ Urban	1	13.82	182	33
IKONOS Forest	1	2.12	41	106
IKONOS Urban	1	8.10	99	81
Aerial Photograph	0.25	30.03	519	91

Table 2. Quantified Image Characteristics (Resolution, Context and contrast)

Correlation analysis provides an objective assessment of the association between pairs of measured variables. A correlation analysis was conducted to examine relationships between these criteria with extraction accuracy quality measures of completeness, correctness and quality (Table 3).

	Completeness	Correctness	Quality
Resolution	0.14	0.52	0.71
Road Density	0.12	-0.75	-0.70
Number of Intersections	0.21	-0.70	-0.65
Contrast	0.71	0.73	0.80

Table 3. Correlation Matrix (Image characteristics vs quality measures)

The correlation analysis showed that good contrast is essential to accurate identification and extraction of roads from imagery. Contrast plays most important role to determine the quality of road extraction than any other criteria. The context characteristics of image (road density and number of intersections) are indications of potential problems for the automatic road extraction.

V. Conclusion/Significant Accomplishments

An integrated approach to automatic road extraction from remotely sensed imagery was successfully developed combining digital image processing and geographic information system technologies. The approach was based on differential geometry while roads were modeled as continuous single lines with bar-shaped or parabolic-shaped profiles. Roads were extracted from the second derivative image and further refined with GIS operations. All the algorithms were developed and integrated in a desktop GIS environment. This software package called Feature Extraction was developed as an extension of ArcView.

Four kinds of commonly used remote sensing data i.e. 30 meter Landsat7 ETM+, 1 meter IKONOS, 1 meter DOQ, and 0.25 meter scanned aerial photography were used to develop this approach. Nine sub-images with rural, suburban, residential and urban context were used. Accuracy assessment of these nine test images demonstrated that the approach was successful. Average completeness across the test imagery was higher than 80%. For rural and suburban areas, the completeness was higher (90%+). The results further identify that the approach is excellent when using images with good contrast. It has great potential for updating and maintaining GIS data at local and regional scale.

The correlation analysis indicated that contrast between road and background greatly affects the quality of the road extraction. Better contrast leads to better results. Image context as measured by road density and number of intersections can also provide indicators. Areas with complex roads will be more difficult to extract road from and lead to low quality measures.

Automatic road extraction from remotely sensed imagery has the potential to save time and money in GIS data collection and update. By improving these approaches and software, it will be very useful within the remote sensing and GIS communities.

VI. Journal/Conference Publication List

An Integrated Approach of Automatic Road Extraction and Evaluation from Remotely Sensed Imagery, *Proceedings of the International Cartographic Conference (ICC2001)*, 6-10 August 2001, Beijing, China,

Automated Feature Extraction from 1-Meter Imagery, *ASPRS Annual Conference*, April 23-27, 2001, St. Louis, MO

Extracting Road Features from Digital Imagery Sources, *Missouri GIS Conference*, March 26-28, 2001, Columbia, MO

VII. Students Supported

Wenbo Song, MA, 2001,

Thesis title: An Integrated Approach of Automatic Road Extraction and Evaluation from Remotely Sensed Imagery

VIII. Subject Inventions

A road extraction software package called 'Feature Extraction' (ArcView extension) was developed with support from this grant.