2009 10th International Conference on Document Analysis and Recognition

Raster Map Image Analysis

Thomas C. Henderson and Trevor Linton
School of Computing
University of Utah
Salt Lake City, UT 84112 USA
tch@cs.utah.edu

Abstract

Raster map images (e.g., USGS) provide much information in digital form; however, the color assignments and pixel labels leave many serious ambiguities. A color histogram classification scheme is described, followed by the application of a tensor voting method to classify linear features in the map as well as intersections in linear feature networks. The major result is an excellent segmentation of roads, and road intersections are detected with about 93% recall and 66 % precision.

1. Introduction

Digital maps contain a wealth of information which can be used for a variety of applications, including the analysis of cultural features, topographical terrain shape, land use classes, transportation networks, or maps can be registered (conflated) with aerial images in order to localize and identify photo imagery structures. Unfortunately, raster map images are typically encoded in such a way that semantic features are difficult to extract due to noise, error or overlapping features. Semantic features of interest include roads, road intersectoins, water regions, vegetation, political boundaries, and iso-elevation contours.

We propose to exploit specific knowledge about the use of color in maps (USGS maps, in particular) as the basis for a color histogram classification first step, and then use this as the starting point for a tensor voting method for the segmentation of linear features (roads, rivers, iso-contours, etc.) and road intersections. The tensor voting method uses a special representation to allow linear pixel level features to reinforce other such features to increase belief in larger scale segments Likewise, pixel-level terminations, corners and crossing points can be precisely determined by input (votes) from neighboring pixels.

More specifically, the following Gestalt principles will be used to segment semantic features; also given is the technical approach for exploiting the principles:

- *Similarity*: Color histograms will be used for all categories, and parallel orientation between linear elements will be used for city street line pairs.
- *Proximity*: Spatial proximity will be emphasized both in the tensor voting method, A*, as well as in the city road line pair analysis. While proximity is important Gestalt's principle of continuity is taken to be more important and emphasized more by the tensor voting method than proximity.
- *Continuation*: Continuation is discovered through the tensor voting method.
- *Closure*: The ability to fill gaps and nd the most optimal curve to close a region. This is done with the tensor voting method.

Analysis of the Gestalt's principle of similarity is performed first with a histogram analysis. Similar segments are identied in the histogram analysis and all similar components are extracted to separate sparse raster maps. Once the sparse raster maps are created they are further treated to clean noise out using dilation and eroding techniques. Smaller features such as parallel road lines or distinct subtle features are improved to prevent loss of these details in the Tensor Voting framework.

These segmentation methods are applied to USGS maps, and the performance is analyzed. For general linear features, a qualitative analysis is given in which a human observer estimates the quality of the segmented features, and the overall result is excellent for the segmentation of roads, rivers and iso-contours. Road intersection detection results are also give; here the performance is analyzed in terms of recall and precision with respect to a set of ground truth road intersections from sub-images in USGS maps. The recall is about 93% while precision is about 66%.

978-0-7695-3725-2/09 $25.00 © 2009 IEEE
DOI 10.1109/ICDAR.2009.31

2. Related Work

Raster map image analysis has a long history [**?**], and more recent work has focused on various aspects of the problems studied here. Many of these methods attempt more general solutions, and do not take advantage of the specific map knowledge as is exploited here. Also, the analysis of linear features in other types of documents (e.g., engineering drawings) has been reported in the literature [**?**, **?**, **?**]. A significant feature of maps is that the road networks, iso-contours, political boundaries, etc. offer a more extensive and coherent linear network than other types of documents; this is also exploited by our approach.

Podlaso et al. [**?**] propose a mathematical morphology approach to the extraction of map classes, but do not exploit the deep knowledge of the use of color in the map as we do here. Chiang et al. have proposed methods for both road intersection detection [**?**] and line pixel classification [**?**] (achieving 75% recall and 90% precision in road detection). Our method provides an improvement in the recall statistic (i.e., the number of correctly extracted intersections divided by the number of ground truth intersections).

3. Color Histogram Pixel Classification

The number of colors in a USGS map is limited to a few specific colors, and this information is exploited in the analysis of the USGA maps. Note that the number of colors used may differ from map to map (from 6 to 13 colors, including black). USGS maps have a well-defined structure which is exploited here to extract semantic contents. The elements of the legend are described in terms of their location in the map image, and their constituent (non-white) pixel values. Note that the pixel values in a USGS image are coded as follows:

Index Value	*Color*	*R*	*G*	*B*
0	Black	0	0	0
1	White	255	255	255
2	Blue	0	151	164
3	Red	203	0	23
4	Brown	131	66	37
5	Green	201	234	157
6	Purple	137	51	128
7	Yellow	255	234	0
8	Light Blue	167	226	226
9	Light Red	255	184	184
10	Light Purple	218	179	214
11	Light Gray	209	209	209
12	Light Brown	207	164	142

From the legend, representative color histograms can be found; for example, consider *Primary Highways* where the histogram is:

Index Value	*Color*	Number in Sub-image
0	Black	1030
1	White	2925
2	Blue	7
3	Red	456
4	Brown	352
5-12	Green	0

This kind of information is exploited to produce a set of rules for classification of image pixels.

The color usage information is determined from the map legend. However, color histogram information alone is not sufficient to discriminate all classes; for example, roads of different types may have the same color histogram and may only be distinguished by their width (in pixels). It is also possible that some classes have erroneous mixtures of color pixels (e.g., waterways may have a significant amount of black, especially near roads). Given the color histogram data extracted from the map legend, the following set of rules were constructed.

```
for each pixel
  Get max_color and max_count
  if (max_color is BLUE) and
     (max_count > BLUE_THRESH)
    then class is WATER
  elseif (max_count is BLACK) and
         (amount of RED < BLUE) and
         (amount of BROWN < BLUE)
    then class is LIGHT_ROAD
  elseif (max_count is BLACK) and
         (amount BLUE < RED) and
         (amount BLUE < BROWN) and
         (amount BLUE > 0) and
         (amount BROWN > 0)
    then class is PRIMARY_ROAD
  elseif (amount BLACK < POLI_BLACK) &
         (amount BLUE < POLI_BLUE) and
         (POLI_RED < amount RED) and
         (POLI_BROWN < amount BROWN)
    then class is POLITICAL BOUNDARY
  elseif (max_color is black) and
         (amount BLUE < BROWN) and
         ((amount RED < BROWN) or
     (BLACK_THRESH < amount of BLACK))
    then class is TRANSPORT
  elseif (max_color is BROWN) and
         (BROWN_THRESH < amount BROWN)
    then class is ISO-CONTOUR
```

An example USGS sub-image is shown in Figure 1 and the results of the application of this method are shown in Figure 2 as binary classification images.

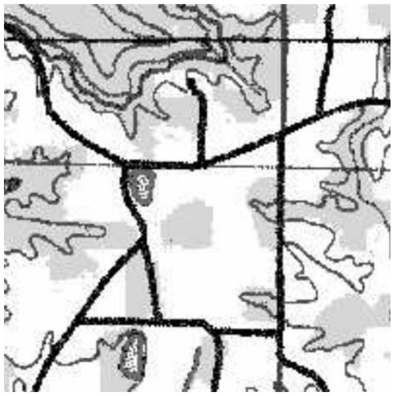

Figure 1. Sub-image from USGS map.

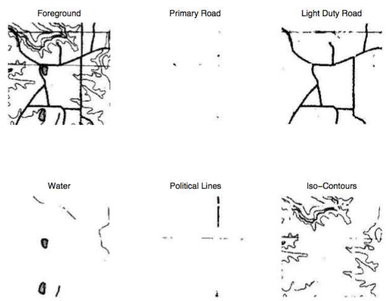

Figure 2. Classification Result for Example Sub-Image.

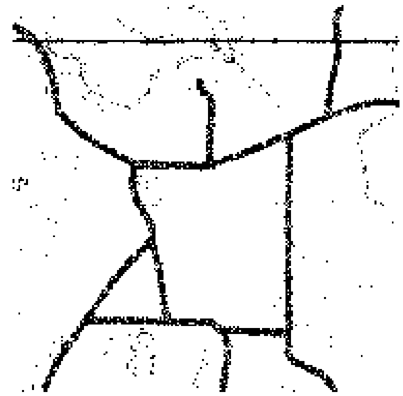

Figure 3. USGS Binarized Sample Map Sub-Image.

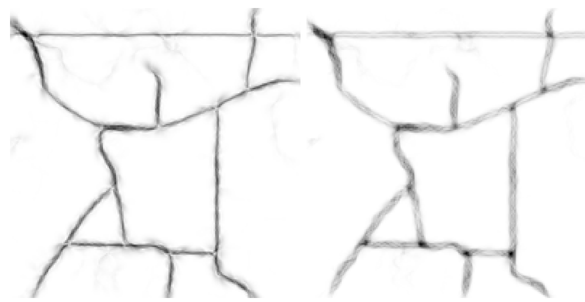

Figure 4. Tensor Voting Images Example Sub-Image.

4. Tensor Voting Method

The sparse independent raster maps are processed by the Tensor Voting framework (see Mordohai and Medioni [?] for complete details). This approach helps remove noise, identifies regions, end points and approximates curves. This process is applied to each raster map but does not necessarily have to be done serially: parallel processing at this point can greatly improve the speed of this framework.

The output of the Tensor Voting framework is a junction map and a curvature map. Figure 3 shows the original sparse raster map before being fed into the Tensor Voting framework. The output featured in Figure 4 shows the resulting curvature and junction maps. Curvature maps express the optimal curves found by the Tensor Voting framework. In this map unit width curves can be extracted through a process of A*, thinning or finding the local maxima across the normal to the path of the curve. The junction map expresses intersections, and though imperfect, it provides a decent guide to finding crossings in the line segments. Other methods must be used to help the extraction of intersections. In addition to curvature and junction maps, examining the raw tensors helps reconstruct regions and find endpoints. The Tensor Voting framework is ideal for our needs. It reinforces many of the Gestalt principles that are critical for feature extraction. When filling gaps the proximity of the closest other line segment is considered however the principle of continuity is also preserved in that the Tensor Voting framework will greatly favor connecting two line segments which are on the same logical curved path or line. Line segments which are close in proximity but have little continuity may still be connected if no other ideal line segment can be found. The following equations show how attenuation fields are created mathematically allowing the creation of a stick tensor voting field and the creation of a ball tensor voting field. The only free variable in the system is σ which controls the size of the voting field. The value of c is defined by Medioni and depends on σ. Small values of c result in very circular attenuation field and favor proximity more than continuity. Large values of c result in a line attenuation field and heavily favor linear continuity.

$$c = \frac{-16 log 0.1(\sigma - 1)}{\pi^2}$$

$$s = \frac{\theta l}{sin(\theta)}$$

$$k = \frac{2sin(\theta)}{l}$$

$$VF_s = (e^{-\frac{s^2+ck^2}{\sigma^2}})gg^T$$

where

$$g = [-sin(2\theta)cos(2\theta)]^T$$

Tensor voting is used to find roads and intersections within the raster map. After the pre-processing histogram analysis the rough estimate of roads in the form of a binary image is fed into the tensor voting system with a σ of 6.5 (29x29 tensor voting field). The resulting curvature and junction maps are then examined for roads and intersections.

A rough estimate of the curvatures and roads are produced from the output of the tensor voting system. The curvature map is put through a process of examining local maxima across the direction of the normal to the tensor. The junction map is examined by looking for regional local maxima.

This method has limitations depending on the size of the roads with-in the USGS map. Smaller compact roads benefit from smaller σ values and vise versa. A larger σ value ruins fine details in smaller town roads while the smaller σ inadequately fills in gaps in longer country roads. The default value used of 6.5 does well with both. A method of using dynamic σ values would be advantageous for future work.

5. Road Segmentation

The tensor voting method described above has been applied to road segmentation in USGS images. This includes various types of roads with difference color histograms (e.g., major roads have a mix of red and black pixels while primary and secondary roads have black, blue and brown pixels). A set of 100 randomly sampled 200x200 subimages from two USGS raster map images were analyzed. The performance was qualitatively evaluated by human observation of each sub-image; the segmentation ratings were based on the following scale: Poor (0-49%), Good (50-69%), Very Good (70-94%), Excellent (95-100%).

The sub-image is selected randomly by the program, then the road segmentation is performed, and finally the segmented roads are overlayed in red on the original subimage and displayed. The human observer then decides the performance rating. An example image result is shown in Figure 5. As can be seen, the road segmentation is excellent.

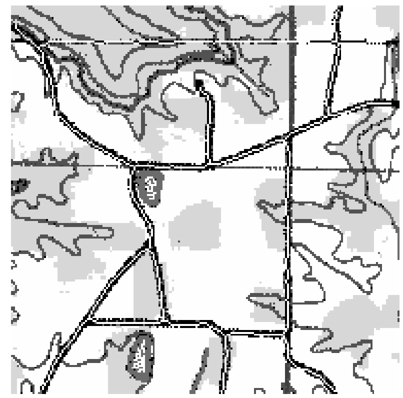

Figure 5. Roads for Test Sub-Image 11 from F34086E1.

6. Road Intersection Detection

Road intersections are particularly useful for the registration of maps to aerial imagery. Typically, registration is performed by determination of the correspondence between a set of points from the map and a set of points from the aeiral image. A transformation function can be defined by the point correspondences combined with geometric constraints on the type of transformation. The tensor voting method described above has been applied to USGS maps.

6.1 Method

A set of 100 randomly selected sub-images from USGS maps has been generated such that at least five road intersections exist in the sub-image. Ground truth road intersections were then determined by a human observer where the (row,column) location was chosen as the intersection of the intersecting roads. Once all the road intersections were determined, the road intersection algorithm was run on the sub-images. A detected intersection matches ground truth if within a threshold distance (five pixels Euclidean distance in this case). The statistics are defined as:

$$Recall = \frac{\mid \{relevant\} \cap \{retrieved\} \mid}{\mid \{relevant\} \mid}$$

$$Precision = \frac{\mid \{relevant\} \cap \{retrieved\} \mid}{\mid \{retrieved\} \mid}$$

With regard to the calculation of the precision statistic, all retrieved road intersections within five pixels of a relevant intersection are counted as one retrieved intersection.

6.2. Results

For the example image in Figure 1, the road intersections are shown in Figure 6.

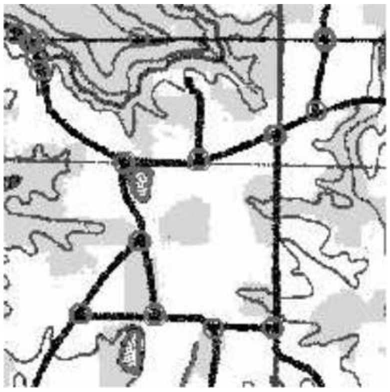

Figure 6. Road Intersections for Example Sub-Image.

The overall result was: 93% recall and 66% precision.

7. Conclusions and Future Work

A preprocessing color histogram method for semantic feature classification as well as a tensor voting method for road segmentation and road intersection detection have been described. The results are qualitatively excellent for road segmentation and produced good results quantitatively for recall and precision on intersections. The computational efficiency using tensor voting is good, a 200 by 200 area took on average 38 seconds to process. It should be possible to improve the road intersection detection precision by more careful analysis of the context of the retrieved intersection. Work is currently in progress on that. It is possible that the road intersection detection recall can be improved, but it seems that this will require the application of broader knowledge-based approach. This is also currently under study.

We would like to acknowledge the support of IAVO Research, Inc. through contract FA9550-08-C-0005.