

Strabo: A System for Extracting Road Vector Data from Raster Maps (Demo Paper)

Yao-Yi Chiang and Craig A. Knoblock

University of Southern California

Department of Computer Science and Information Sciences Institute

4676 Admiralty Way, Marina del Rey, CA 90292, USA

[yaoyichi, knoblock]@isi.edu

ABSTRACT

Raster maps contain valuable road information, which is especially important for the areas where road vector data are otherwise not readily accessible. However, converting the road information in raster maps to road vector data usually requires significant user effort to achieve high accuracy. In this demo, we present Strabo, which is a system that extracts road vector data from heterogeneous raster maps. We demonstrate Strabo's fully automatic technique for extracting road vector data from raster maps with good image quality and the semi-automatic technique for handling raster maps with poor image quality. We show that Strabo requires minimal user input for extracting road vector data from raster maps with varying map complexity (i.e., overlapping features in maps) and image quality.

Categories and Subject Descriptors

H.2.8 [Database Management]: Database Applications—*Spatial Databases and GIS*

General Terms

Algorithms, Design

Keywords

GIS, raster maps, road vectorization, map processing

1. INTRODUCTION

Raster maps are easily accessible than other geospatial data, such as imagery, and provide geographic features that are difficult to find elsewhere, such as landmarks in historical maps. Since roads commonly exist across various geospatial data (e.g., satellite imagery), by extracting the road vector data from a raster map and matching the road geometry with the road geometry from a georeferenced data set, we can identify the geospatial extent of the raster map and align the raster map with other georeferenced data [1, 5]. Further, we can provide road vector data for the areas where

such data are otherwise not available. Extracting accurate road vector data from raster maps is challenging because of the varying image quality, the complexity of maps (i.e., overlapping features in maps), and the typical lack of metadata (e.g., map geocoordinates, map source, original vector data). To overcome these difficulties, in this demo, we present Strabo, which is a general road-vectorization system for processing heterogeneous raster maps with minimal user input [2, 3, 5]. For raster maps with good image quality, Strabo automatically extracts the foreground pixels and then generates the road geometry and the road vector data [2, 5]. For complex raster maps, or maps with poor image quality, Strabo employs a semi-automatic technique that requires a few examples of road areas to first extract the road pixels from the maps [3]. Then, Strabo automatically produces the road vector data from the extracted road pixels [2].

2. RELATED WORK

The research for vectorizing raster maps in the earlier years [7] was highly labor-intensive and the research focused on hardware design and the techniques to help human operators to speed up the whole processes. For example, Leberl and Olson [7] describe techniques to aid the operator to move the cursor quickly, such as the line-following technique to automate the cursor movement. In recent work, Itonaga et al. [6] work on non-scanned maps to extract the road vector data with manually specified parameters, such as the road width. In comparison, Strabo handles heterogeneous map types and extracts road vector data from raster maps with minimal user input using dynamically generated parameters.

In addition to the research work, a commercial product called R2V from Able Software is an automated raster-to-vector conversion software package specialized for digitizing raster maps. R2V requires significant user effort on labeling road areas and parameter tuning to achieve accurate results compared to Strabo, especially for maps with numerous colors, such as scanned maps [2].

3. STRABO ARCHITECTURE

Strabo consists of two major components: a fully automatic and a semi-automatic road vectorization components. To vectorize the roads in a raster map, the user first invokes the fully automatic road vectorization function of Strabo. The user then reviews the vectorization results and decides if the semi-automatic approach is required to improve the results. We briefly explain the key ideas of the automatic and semi-automatic components in the following subsections.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ACM GIS '10 San Jose, CA USA

Copyright 2010 ACM 0-12345-67-8/90/01 ...\$10.00.

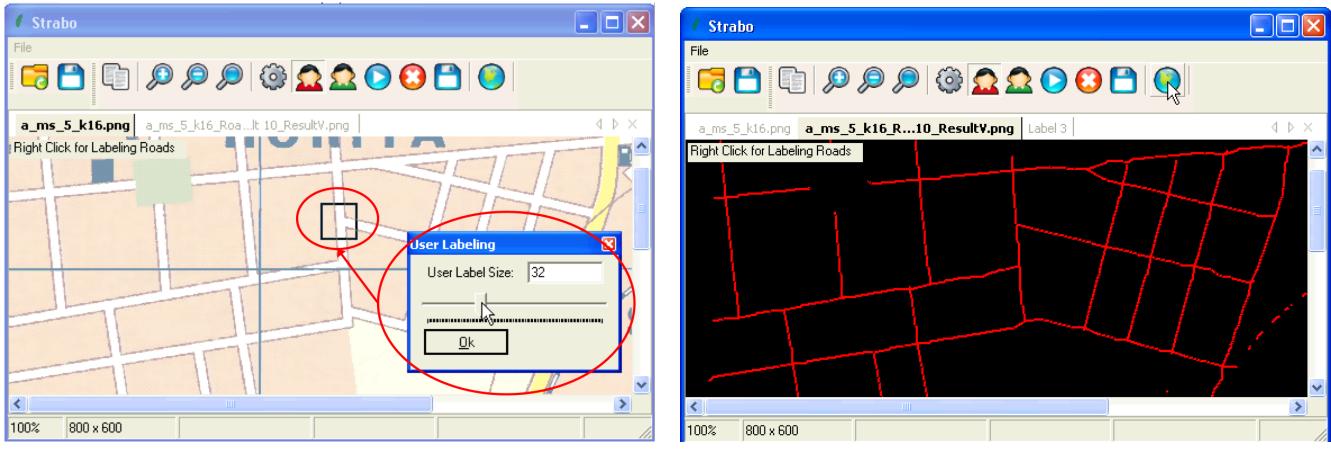


Figure 1: Snapshots of Strabo

3.1 Automatic Road Vectorization

First, to identify the road pixels from raster maps automatically, Strabo utilizes the heuristic that the background colors of raster maps have a dominant number of pixels and the foreground colors have high contrast against the background colors to remove the background pixels. Next, to separate the overlapping features in the foreground pixels, the automatic technique exploits the distinctive geometric properties of the road lines to detect the road width and road format (i.e., single line or double lines) and identifies the road pixels with dynamically generated parameters, such as a road line should be longer than the detected road width. Finally, Strabo links the identified road pixels to extract and vectorize the centerlines of the linked road pixels for producing the road vector data.

3.2 Semi-Automatic Road Vectorization

The automatic road vectorization approach of Strabo does not work well for maps with poor image quality since maps with poor image quality generally contain numerous colors and hence the background colors do not always have a dominant number of pixels. Therefore, Strabo provides a semi-automatic technique with a supervised user-labeling function to first extract the road pixels [3].

To start the user-labeling function, Strabo first generates a set of quantized maps with reduced numbers of colors. The user then selects a quantized image that contains road lines shown in different colors compared with other map features in order to label the road areas. The goal of the user-labeling task is to provide a user label for each road color in the quantized map. For example, for the quantized map shown in Figure 1, the user needs to provide two user labels, one for the white roads and one for the yellow roads. A user label is a rectangle that should be large enough to cover a road intersection or a road segment. The user-labeling function of Strabo provides a dialog interface with a slide bar for the user to indicate the size of a user label, as shown in Figure 1(a). The user label should be centered at a road intersection or at the center of a road line (approximately), which is the property that Strabo exploits to identify the road colors from the user labels automatically. After the user-labeling task is finished, Strabo analyzes the user labels to extract the road pixels and then generates the road vector data from the extracted road pixels. Figure 1(b) shows the

extracted road vector data from the raster map shown in Figure 1(a).

4. DISCUSSION AND FUTURE WORK

We presented Strabo, which is a system that requires minimal user effort for extracting accurate road vector data from heterogeneous raster maps. The average processing time for Strabo to vectorize the road pixels for an 800x550-pixels map was 5 seconds and for a 2084x2756-pixels map was 2 minutes. In the future, we plan to incorporate into Strabo the functionality for recognizing text labels in raster maps from our previous work [4].

5. ACKNOWLEDGMENTS

This research is based upon work supported in part by the University of Southern California under the Viterbi School Doctoral Fellowship.

References

- [1] Chen, C.-C., Knoblock, C. A., and Shahabi, C. (2008). Automatically and accurately conflation raster maps with orthoimagery. *GeoInformatica*, 12(3):377–410.
- [2] Chiang, Y.-Y. and Knoblock, C. A. (2009a). Extracting road vector data from raster maps. In GREC, LNCS, 6020, pages 93–105, New York. Springer.
- [3] Chiang, Y.-Y. and Knoblock, C. A. (2009b). A method for automatically extracting road layers from raster maps. In *Proceedings of the Tenth ICDAR*.
- [4] Chiang, Y.-Y. and Knoblock, C. A. (2010). An approach for recognizing text labels in raster maps. In *Proceedings of the 20th ICPR*.
- [5] Chiang, Y.-Y., Knoblock, C. A., Shahabi, C., and Chen, C.-C. (2008). Automatic and accurate extraction of road intersections from raster maps. *GeoInformatica*, 13(2):121–157.
- [6] Itonaga, W., Matsuda, I., Yoneyama, N., and Ito, S. (2003). Automatic extraction of road networks from map images. *Electronics and Communications in Japan (Part II: Electronics)*, 86(4):62–72.

- [7] Leberl, F. W. and Olson, D. (1982). Raster scanning for operational digitizing of graphical data. *Photogrammetric Engineering and Remote Sensing*, 48(4):615–672.