

PedSeg: GPS Tracks as Priors for Overhead Image Segmentation (Demo Paper)

Shawn Newsam, Brent Edmunds, and Andrew Pierce
University of California at Merced
snewsam,bedmunds,apierce2@ucmerced.edu

ABSTRACT

We introduce PedSeg, a system for computing the boundaries of visually distinct geo-spatial objects. Knowing the precise boundaries or even bounding box approximations of geo-spatial objects is important for maintaining existing repositories such as gazetteers as well creating volunteered geographic information sources such as OpenStreetMap. PedSeg uses active contour image segmentation to determine an object's spatial extent from high-resolution overhead imagery. The novel aspect of this work is that the image segmentation is seeded with a GPS track acquired by simply walking around or otherwise traversing the approximate boundary of the target object (thus the prefix Ped). The technique is intended to be completely automated once the GPS track has been loaded into the system. This provides several advantages such as the user not needing to be skilled at image editing. In fact, the user does not even need to access, view, or interact with the imagery. The use of active contour image segmentation compensates for inaccuracies in the user acquired GPS tracks due to GPS error or the physical inaccessibility of the object boundary. We present examples of using PedSeg to compute the boundaries of a number of geo-spatial objects.

Categories and Subject Descriptors

H.2.8 [Database Management]: Database Applications—*spatial databases and GIS*; I.5.4 [Pattern Recognition]: Applications; I.4.8 [Image Processing and Computer Vision]: Scene Analysis

Keywords

image segmentation

1. INTRODUCTION

This paper describes PedSeg, a system for computing the boundaries of visually distinct geo-spatial objects. At the core of the system is active contour image segmentation for

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 SIGSPATIAL GIS '11, November 1-4, 2011. Chicago, IL, USA
Copyright ©2011 ACM ISBN 978-1-4503-1031-4/11/11 ...\$10.00.

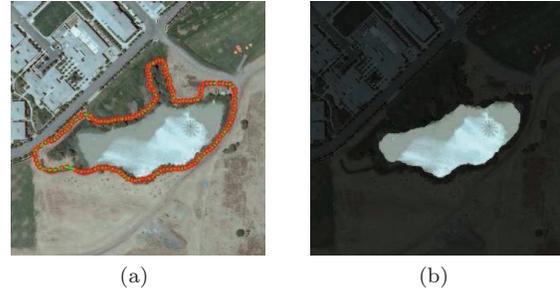


Figure 1: (a) A GPS track obtained by walking around a small lake on the UC Merced campus. This is the input to the PedSeg system. (b) The boundary as computed by the PedSeg system.

automatically determining an object's spatial extent from high-resolution overhead imagery. The novelty of the system lies in the use of GPS tracks as priors for the segmentation. In particular, user acquired GPS tracks are used to seed the segmentation algorithms. The technique is intended to be completely automated once the GPS track has been loaded into the system. This provides several advantages such as the user not needing to be skilled at image editing. In fact, the user does not even need to access, view, or interact with the imagery.

Knowing the precise boundaries or even bounding box approximations of geo-spatial objects is important for a number of geographical information systems (GIS) and their applications. The United States Geological Survey (USGS) Geographic Names Information System (GNIS)¹ and similar gazetteers are becoming increasingly useful as a source of what-is-where on the surface of the Earth. However, these gazetteers have shortcomings, some of which could potentially be overcome by advances in automated image analysis. In particular, the spatial extent of the records is currently limited to a single point, a longitude/latitude pair. While gazetteers typically include provisions for storing at least a bounding box representation, this information has simply never been acquired. As the University of California at Santa Barbara Alexandria Digital Library (ADL) gazetteer development team pointed out [4], “for a digital library application, the spatial extent of the feature, either approximately with a bounding box or more accurately with a polygonal representation, is better, but there are no large sets of gazetteer data with spatial extents.” They go on

¹<http://geonames.usgs.gov>

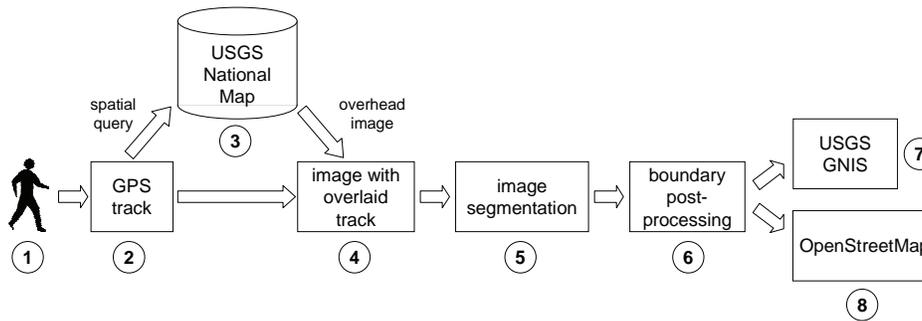


Figure 2: PedSeg system architecture.

to state that, “establishing the standards that will enable the sharing of gazetteer data will help harvest data from many sources, but ultimately deriving spatial locations and extents from digital mapping products and other sources automatically will be needed.” The PedSeg system innovatively combines two digital sources—GPS tracks and high-resolution overhead imagery—to derive the spatial locations and extents for a number of geo-spatial object types.

Determining the spatial extent of geo-spatial objects is also fundamental to volunteered geographic information (VGI) projects such as OpenStreetMap², a “free editable map of the whole world. . . made by people like you.” One immediate advantage that OpenStreetMap maps have over the copyrighted commercial maps though is that they can be preloaded or cached by navigation applications and can thus be used in remote locations on portable devices that do not have cell-phone connections or are not able to receive service. OpenStreetMaps records are commonly created with the help of user-provided GPS tracks. However, the tracks only provide an approximation of an object’s true boundary so the user has to manually edit the data typically using an image underlay. The intent of the PedSeg system is to acquire the boundaries without requiring the user to have direct access to or even see or interact with the overhead imagery. All the user needs to do is physically traverse the object of interest with a GPS logger and then upload the track. This allows a greater number of users, particularly those that are not computer savvy, to contribute to OpenStreetMap and similar VGI projects.

2. SYSTEM OVERVIEW

A diagram of the PedSeg system architecture is shown in figure 2. The various components will be described in more detail below but the process of using a user-acquired GPS track to determine the boundary of a geo-spatial object is as follows. First, the user walks around or otherwise traverses the boundary of an object while using a GPS logger to record the track information. The track is then input to the system and the locations of the track points are used to issue a spatial query to the USGS National Map to retrieve co-located high-resolution overhead imagery. The track is then registered and overlaid on the imagery and passed to the image segmentation module. The segmentation algorithm uses the track as the initial boundary and then refines the boundary based on the pixel values of the image. Finally, the object boundary is post-processed and can be fed

²<http://www.openstreetmap.org>

to gazetteers such as the USGS GNIS, VGI projects such as OpenStreetMap, or other repositories.

3. SYSTEM DETAILS

3.1 Querying The National Map

PedSeg uses the National Map Seamless Data Server³ interface to automatically download high-resolution overhead imagery. This interface accepts spatial queries for a range of data collections including High Resolution Orthoimagery (HRO) of major U.S. urban areas at 3-inch, 6-inch, 1-foot, and 2.5-foot spatial resolutions, and the U.S. Department of Agriculture (USDA) National Agriculture Imagery Program (NAIP) imagery of the conterminous United States at 1-meter or 2-meter spatial resolutions.

Spatial queries to the National Map require both an address and geographic coordinates. PedSeg uses Google’s reverse geocoding service⁴ to determine the approximate address of the track based on the latitude and longitude (in World Geodetic System 84) of the first track point. The image download then takes place in three steps. First, the reverse-geocoded address and a slightly enlarged bounding box of the GPS track points are used to query the National Map to determine what image data is available for the location. If HRO (preferred) or NAIP imagery is available then a data request query is issued. A URL is then generated which PedSeg uses to download the imagery.

3.2 Overlaying the GPS Tracks on the Images

The download set from the National Map Seamless Data Server includes meta-data indicating the spatial extent of the image. This allows us to determine the locations of the GPS track points in the image’s reference frame. The image and the pixel coordinates of the track points are then passed to the segmentation module.

3.3 Active Contour Image Segmentation

The intent of the PedSeg system is to extract an object’s boundary without the user seeing or interacting with the overhead imagery. We therefore chose active contour segmentation over interactive methods such as intelligent scissors [5] or graph cut [1]. We also chose active contour segmentation over other methods which also only require an initial boundary such as GrabCut [7] since active contours can be seeded with tracks that are completely inside the boundary of the target object.

³<http://seamless.usgs.gov>

⁴<http://code.google.com/apis/maps/documentation/geocoding>

Active contours, at their core, are methods that evolve an initial curve according to some minimization criteria. Ideally, the curve will identify and segment all of the objects in an image. Unfortunately, when run on large multifeatured images like those from a satellite, active contours do not fare well—it takes rather specific initial curves to identify a desired object. We overcome this issue by using the GPS tracks to initialize our curve. Our a priori assumption is that the GPS tracks roughly approximate the desired object and thus are an appropriate seed for the contour. The GPS track points are turned into a curve through a naive (linear) approximation. This is not necessarily the most effective method, as the GPS tracks can occasionally be off. In future work we will investigate a more adaptive interpolation scheme, for instance B-splines.

A popular technique in curve evolution, introduced by Osher and Sethian [6], is mean curvature. Mean curvature, when used with level sets, allows the curve to avoid local minima, and automatically alter topology. Also, the problem can be discretized by a fixed rectangular grid.

We considered two extensions of the mean curvature method. First, the Chan-Vese Model (C-V model) [3] which searches for the smallest curve that separates areas of differing intensity. Unfortunately, as most real world objects are made of regions of varying intensity, this did not result in the segmentation we desire. For instance, the Quad image in figure 6, when segmented using the C-V model, will only return the grass, ignoring the sidewalks. This type of result is not useful for our purposes, as we wish for a tight boundary.

We therefore settled on geodesic active contours [2], which minimize the following functional:

$$\inf_c J(C) = \int_0^1 G(\nabla u(C(s))) \cdot |C'(s)| ds \quad (1)$$

where u is the input image, ∇u is an edge detector, C is the contour, $G(x)$ is some function that as $x \rightarrow \infty$, $G(x) \rightarrow 0$. The minimization problem (1) is solved when $G(\nabla u(C(s)))$ approaches zero—i.e., when the curve C approaches a boundary. The function $G(\nabla u(C(s)))$ when multiplied by $|C'(s)|$ finds the smallest curve with the largest gradient.

3.4 Boundary Post-Processing

The complexity of natural images often results in segmentation artifacts such as holes, thin protrusions, and small gulfs. We apply standard post-processing techniques such as image infilling, morphological opening and closing, and convex hull computation to improve the final boundary.

4. SAMPLE RESULTS

Figures 3 through 7 show the results of applying the Ped-Seg system to GPS tracks acquired by walking around a number of objects in different geographic locations. These tracks were acquired using a Visiontac Multifunction GPS Data Logger (Model VGPS-900), a low-cost consumer grade logger which retails for approximately 100 US dollars. The objects are as follows: a small lake on the UC Merced Campus; a fountain in San Francisco; a sports track in Santa Barbara; a quad on the UC Merced campus; and a small pond also on the UC Merced campus. The images for the lake, quad, and pond have a resolution of 1m and the images for the fountain and sports track have a resolution of 1ft.

For each result, we show (a) the image as retrieved from the National Map; (b) the GPS track overlaid on the image;

(c) the initial contour as derived from the track; (d) the output of the segmentation module; (e) the post-processed boundary; and (f) the final boundary overlaid on the image. The post-processing step varies slightly for each result. We apply infilling to all the segmentation results to remove holes. Morphological opening is applied to the fountain, lake, and pool results based on the heuristic that the segmentation produces an irregularly shaped boundary. Morphological closing is applied to the quad result based on the heuristic that the segmentation produces a rectangular shaped boundary. And, morphological opening followed by a convex hull transformation is applied to the sports track image based on the heuristic that the initial GPS track is itself convex. A rule based system is being implemented to apply these heuristics. Note, however, that the segmentation results tend to be quite good and so post-processing beyond simple infilling is not critical.

These results clearly demonstrate how the active contour segmentation is able to compensate for inaccuracies in the GPS tracks due to GPS error or physical inaccessibility of the object boundary. In the cases of the fountain and sports track, the user did actually return to the starting point of the track but this was not detected by the logger. And, in the case of the lake, the subject was not able to follow the water's edge due to thick foliage. Nevertheless, the segmentation module was able to refine these noisy tracks and detect the true object boundaries.

5. ACKNOWLEDGEMENTS

This work was funded in part by NSF grant IIS-0917069 and a Department of Energy Early Career Scientist and Engineer/PECASE award. Any opinions, findings, and conclusions or recommendations expressed in this work are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. The authors would like to thank the anonymous reviewers for their helpful comments.

6. REFERENCES

- [1] Y. Boykov and M.-P. Jolly. Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. In *IEEE International Conference on Computer Vision*, volume 1, pages 105–112, 2001.
- [2] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *International Journal of Computer Vision*, 22:61–79, February 1997.
- [3] T. F. Chan and L. A. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277, 2001.
- [4] L. L. Hill, J. Frew, and Q. Zheng. Geographic names: The implementation of a gazetteer in a georeferenced digital library. *D-Lib*, 5(1), 1999.
- [5] E. N. Mortensen and W. A. Barrett. Intelligent scissors for image composition. In *ACM SIGGRAPH*, pages 191–198, 1995.
- [6] S. Osher and J. A. Sethian. Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics*, 79(1):12–49, 1988.
- [7] C. Rother, V. Kolmogorov, and A. Blake. “GrabCut”: Interactive foreground extraction using iterated graph cuts. In *ACM SIGGRAPH*, pages 309–314, 2004.

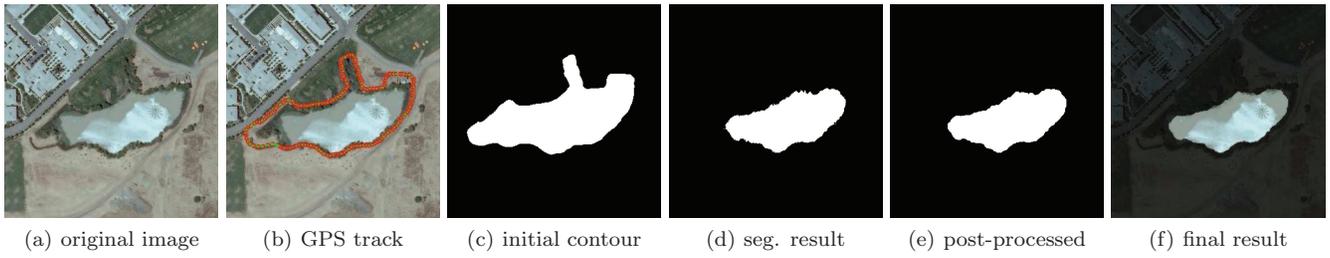


Figure 3: Lake.

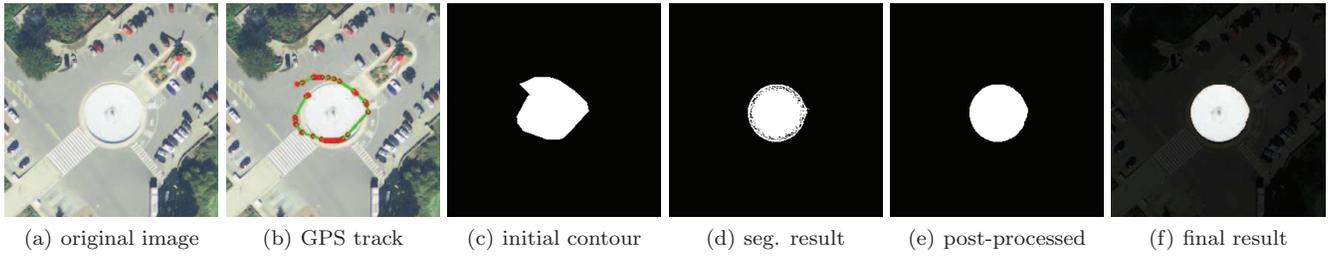


Figure 4: Fountain.

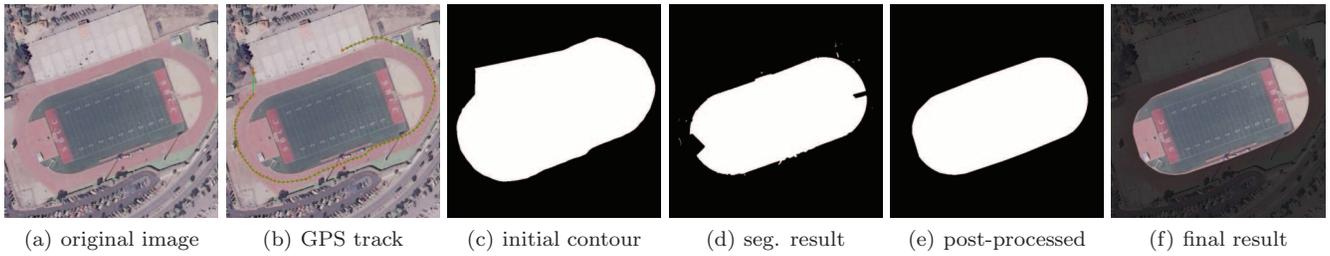


Figure 5: Track.

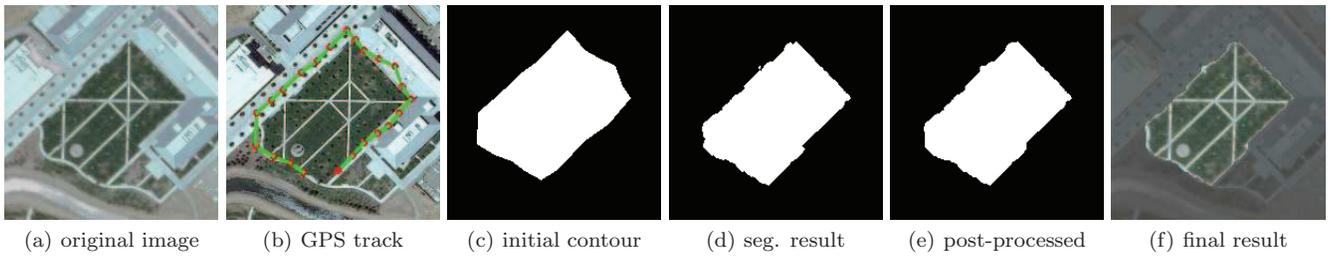


Figure 6: Quad.

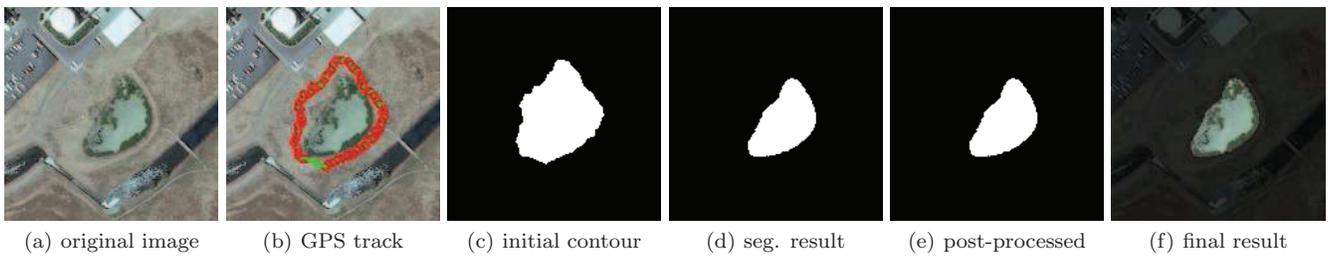


Figure 7: Pond.