# Mobile address tagging for OpenStreetMap

Joel Nelson <jjnelson@stanford.edu> Dave DeBaun <ddebaun@stanford.edu>
Steven Bell <sebell@stanford.edu>

April 25, 2012

Uses DROID smartphone: **Yes**

The objective of this project is to create an Android smartphone application which can tag features in OpenStreetMap (OSM) using the phone's camera and GPS. OpenStreetMap is a crowd-sourced mapping project, analogous to Wikipedia for cartography. While many features such as roads and buildings can be readily drawn from aerial and satellite imagery, a complete map requires that these objects be tagged with their names, addresses, and other data. This has traditionally been a tedious task requiring field notes and hours entering tags on a computer, but mobile devices have the potential to radically accelerate this process.

Our application will allow the user to snap a photograph of a building's address or a business' sign, and text from the image will be detected, parsed and applied as tags to the appropriate object in OpenStreetMap. At a minimum, our application will correctly identify the address of the building in question. We also intend to extract the name, phone number, and website where this information is present.

To tag a building, the application locates the text within the image and parses it to obtain individual tags. Then the application determines the user's location and downloads the OpenStreetMap data for the area immediately surrounding the user. This data is parsed to find the most likely building, the tags are applied, and the changes are uploaded to the OSM server.

Many existing text location methods use the same basic approach. First, the image is segmented based on the fact that text is typically placed against a uniform contrasting background. The set of image regions are then filtered to remove non-letters using a wide range of criteria. Finally, character bounding boxes are merged together to form complete words or sentences. Our application will handle the additional complication that address numbers on houses are often placed vertically or even diagonally. We will test our text detection algorithm using the ICDAR 2003 dataset, which is the most common benchmark for text detection in natural scenes. Our complete algorithm will be tested against our own set of geotagged pictures.

We expect to use an off-the-shelf OCR engine to read the text. Several prior text detection and recognition projects have used Tesseract, but we will also evaluate the performance of Ocrad and CuneiForm.

## References

[1] E. Kim, S. Lee, and J. Kim, "Scene Text Extraction Using Focus of Mobile Camera," *2009 10th International Conference on Document Analysis and Recognition*, pp. 166–170, 2009.

[2] Y. Zhang, K. K. Tan, and T. H. Lee, "Text extraction from images captured via mobile and digital devices," in *2009 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*. IEEE, Jul. 2009, pp. 566–571.

[3] A. Shahab, F. Shafait, and A. Dengel, "ICDAR 2011 Robust Reading Competition Challenge 2: Reading Text in Scene Images," *2011 International Conference on Document Analysis and Recognition*, pp. 1491–1496, Sep. 2011.

[4] H. Chen, S. S. Tsai, G. Schroth, D. M. Chen, R. Grzeszczuk, and B. Girod, "Robust text detection in natural images with edge-enhanced Maximally Stable Extremal Regions," in *Image Processing (ICIP), 2011 18th IEEE International Conference on*, 2011, pp. 2609–2612.

[5] B. Epshtein, E. Ofek, and Y. Wexler, "Detecting text in natural scenes with stroke width transform," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, Jun. 2010, pp. 2963–2970.

[6] S. S. Tsai, H. Chen, D. Chen, R. Vedantham, R. Grzeszczuk, and B. Girod, "Mobile Visual Search Using Image and Text Features," in *Proc. 45th Annual Asilomar Conference on Signals, Systems and Computers*.

[7] S. M. Hanif, L. Prevost, and P. A. Negri, "A cascade detector for text detection in natural scene images," in *2008 19th International Conference on Pattern Recognition*. IEEE, Dec. 2008, pp. 1–4.

[8] a.L. Yuille, "Detecting and reading text in natural scenes," *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 366–373, 2004.