Extraction of road networks from aerial photos with minimum cost paths and P^N-Potts potentials

Jan Dirk Wegner, Javier A. Montoya-Zegarra, and Konrad Schindler

Photogrammetry and Remote Sensing, ETH Zürich, Stefano-Frascini-Platz 5, CH-8093 Zürich (jan.wegner@geod.baug.ethz.ch)

We present a probabilistic representation of network structures in images. Our target application is the extraction of urban roads from aerial images. Roads appear as thin, elongated, partially curved structures forming a loopy graph, and this complex layout requires a prior that goes beyond standard smoothness and co-occurrence assumptions. In the proposed model the network is represented as a union of 1D paths connecting distant (super-)pixels. A large set of putative candidate paths is constructed in such a way that they include the true network as much as possible, by searching for minimum cost paths in the foreground (*road*) likelihood. Selecting the optimal subset of candidate paths is posed as MAP inference in a higher-order conditional random field. Each path forms a higher-order clique with a type of clique potential, which attracts the member nodes of cliques with high cumulative road evidence to the foreground label. That formulation induces a robust P^N-Potts model (Kohli et al., 2009), for which a global MAP solution can be found efficiently with graph cuts. Experiments with two road data sets show that the proposed model significantly improves per-pixel accuracies as well as the overall topological network quality with respect to several baselines.

In our approach the road network is thought of as the union of many elongated *paths*. In this way, network extraction can be cast as the search for a set of paths that together cover the entire network. The proposed method follows the recover-and-select strategy:

> In the *recover* step a large, over-complete set of potential *candidate paths* is generated, by finding the most road-like connections between many different pairs of seed points. The aim of candidate generation is high recall, ideally the candidate set covers the entire road network, at the cost of also containing many false positives that do not lie (completely) on roads.
>
> In the *select* step undesired false positives are pruned from the candidate set to yield a reduced set still covering as much as possible of the network, but with few false positives. This second step is formulated as the minimization of a global higher-order CRF energy, and can be solved to global optimality.

In more detail, our system consists of the following steps: First, an image is segmented into superpixels, which are from then on treated as the smallest entities (nodes) to be labeled. Per superpixel a feature vector is extracted and fed to a binary Random Forest classifier, which assigns each superpixel a unary road likelihood. Next, promising candidate paths are generated. To that end, superpixels with high road likelihoods are sampled randomly as seed nodes and

linked with minimum cost paths. The hope is that *road* superpixels that have high *background* probability, e.g. due to a cast shadow, will be covered by a minimum cost path and thus become member of a clique where the majority of superpixels votes for *road*. The superpixels of each candidate path form a higher-order clique in a CRF. The potentials of these higher-order cliques are based on the P^N-Potts model of Kohli et al. (2009) that enforces *label consistency* within large cliques, meaning that superpixels within the clique are penalized for deviating from the majority label. In that sense, our method could be seen as an anisotropic “smoothing along the paths”. The resulting CRF energy can be minimized with a graph cut, leading to a global optimum of the binary labeling task. Note that working with the actual long-range paths (cliques) is conceptually different from methods that divide long paths into short segments and classify each segment separately (Türetken et al., 2012). Like (Wang et al., 2011; Türetken et al., 2011) we prefer to work with complete paths, so as not to lose any connectivity information.

REFERENCES

Kohli, P., Ladicky, L., Torr, P.H.S., 2008. Robust higher order potentials for enforcing label consistency, in: IEEE Conference on Computer Vision and Pattern Recognition.

Türetken, E., González, G., Blum, C., Fua, P., 2011. Automated Reconstruction of Dendritic and Axonal Trees by Global Optimization with Geometric Priors. Neuroinformatics 9, 279 – 302.

Türetken, E., Benmansour, F., Fua, P., 2012. Automated Reconstruction of Tree Structures using Path Classifiers and Mixed Integer Programming, in: IEEE Conference on Computer Vision and Pattern Recognition.

Wang, Y., Narayanaswamy, A., Roysam, B., 2011. Novel 4-D open-curve active contour and curve completion approach for automated tree structure extraction, in: IEEE Conference on Computer Vision and Pattern Recognition.