SEMANTIC SEGMENTATION OF SETTLEMENT PATTERNS IN GRAY-SCALE MAP IMAGES USING RF AND CRF WITHIN AN HPC ENVIRONMENT

D. Schemala [a], D. Schlesinger [a], P. Winkler [b], H. Herold [c, *], G. Meinel [c]

[a] Chair of Computer Vision, Institute of Artificial Intelligence, TU Dresden, Germany –
dmytro.shlezinger@tu-dresden.de
[b] ZIH - Center for Information Services and High Performance Computing TU Dresden, Germany –
peter.winkler1@tu-dresden.de
[c] Leibniz Institute of Ecological Urban and Regional Development, Dresden, Germany –
h.herold@ioer.de

KEY WORDS: Segmentation, Maps, Machine Learning, Random Forest, Conditional Random Field, High Performance Computing

ABSTRACT:

The segmentation of complex and compound geographical objects such as settlement areas in images is a challenging task. This is particularly true if the images are scanned historical maps, which are only available as gray-scale images. On the other hand, their vast availability and spatio-temporal coverage make these maps a valuable source of information for historical land use and land cover (LULC) research and urban geography. To facilitate the laborious information extraction from these maps, we present a two-stage machine learning-based approach for segmenting settlement and non-settlement areas in the map scans. We employ a Conditional Random Field (CRF) which obtains its unary potentials from a Random Forest (RF). The method is tested using two inference algorithms. To evaluate the performance and the scalability of the approach over large amounts of data sets, we conduct parallel computing experiments within a High Performance Computing (HPC) environment at the Center for Information Services and High Performance Computing at TU Dresden. Experimental results indicate the suitability of both the methodological approach as well as its parallel implementation.

1. INTRODUCTION

Semantic segmentation of complex geographical objects such as settlements in historical map images is a challenging task. However, the historical information is of great interest, e.g. for land use and land cover change (LULCC) research, urban geography or archaeology. In order to facilitate the laborious manual information extraction from historical maps, research on automated map interpretation and cartographic pattern recognition has a long-standing tradition (cf. Herold et al., 2013). While existing approaches to map interpretation are mostly based on extraction of single object types from colour map images, here we follow a different approach. We develop and present a methodology to segment settlement areas as a whole from gray-scale map scans using only a small set of training samples. We evaluated the methodology on a specific map type (in German so-called Messtischblaetter) of large spatial and temporal coverage, which makes them particularly valuable for land change research and historical geoinformation systems (HGIS).

2. METHODOLOGY

The supervised segmentation approach comprises two stages combining a pixel-wise classification and a graphical model, which are described in the following section. Subsequently, in section 2.2 an approach and the technical environment to scale the method to large amounts of input images is described.

2.1 Semantic Segmentation

For pixel-wise segmentation we use a Random Forest (Breiman, 2001). A Random Forest consists of a number of Decision Trees (DT) which, when applied to a pixel, yield the estimated probability that this pixel belongs to settlement area. The outcome of all Decision Trees is added up and normalized to obtain the total probability estimation of the RF. A DT makes its decisions based on the gray levels of the input pixel and of its surrounding pixels. Every test function checks whether or not the difference between two neighbouring pixels is smaller than a threshold. The offsets of these pixels and the threshold are learned during the training phase. Every DT in the Forest is trained individually using an equal number of settlement and non-settlement pixels which are sampled from the training images. The depth of a DT is restricted to a maximum depth to reduce overfitting. We used the information gain as splitting criterion. Since we know a-priori that settlements usually cover largish, connected areas in the image, we smooth the outcome of the RF using a CRF (Lafferty et al., 2001).

The unary potentials for each output variable are simply the probability assigned to by the RF. The pairwise potentials between neighbouring output variables are determined by a constant function which yields a cost e for adjacent variables with different labelling and zero otherwise. A good value of the constant e as well as the various parameters for building the RF are learned using cross validation. As inference algorithm we used the Min-Cut/Max-Flow implementation in (Boykov et al., 2004) for Maximum A-Posteriori Inference. We also applied Maximum-Marginal Inference using Gibbs Sampling.

---

*   Corresponding author

2.2 Parallelization

To evaluate the scalability of the approach for the large and ever increasing amount of available map scans, we perform parallelization experiments on the high performance computing (HPC) cluster TAURUS cluster located at the Center for Information Services and High Performance Computing at TU Dresden. The Taurus architecture has 7,184 CPU cores and a peak performance of 137 TFLOPS.

As the input data comprise a large number of individual images, the appropriate way of parallelization is to perform several hundred independent parallel segmentation jobs.

3. RESULTS

For the evaluation of both the segmentation performance as well as the scalability various experiments have been conducted. Section 3.1 describes the experimental setup and Section 3.2 and 3.3 the results for both the segmentation quality and the parallelization performance.

3.1 Experiments

We divided the data set into training data, test data and validation data. The following experiments have been conducted: We measured the influence of the parameters t (number of trees in the RF), d (the maximum depth of a tree), w (width of the window around each pixel), p (number of samples while building a decision function), e and N (the number of steps when doing Gibbs Sampling) on the segmentation performance. Then, we measured the performance when using different amounts of training data. Furthermore, we tested the generalization of the approach, since the used maps come in several types, each with a slightly different mark-up of settlement area and other elements. We measured the outcome when using test data of the same type as the training data and of a different type. The performance measures used was *Precision* and *Recall*, combined into the F-measure.

3.2 Segmentation

Figure 1 depicts how the F-measure depends on the number of training pixels (cf. Schemala 2016). For the same map image type, the number of necessary training samples is low. Figure 2 shows the F-measure against the pairwise cost e (cf. Schemala 2016). Thus, e is a very crucial parameter to tune.
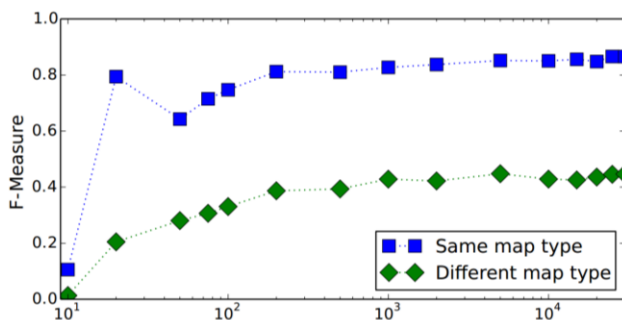

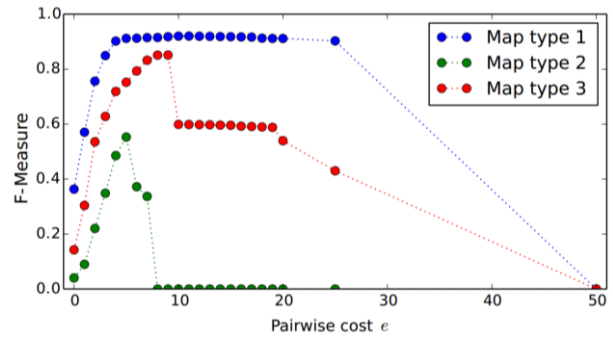Figure 1. The F-measure plotted against the number of pixels used for the training.

Figure 2. F-measure plotted against the pairwise cost e

3.3 Parallelization

Segmentation of 391 maps with a total storage volume of 45 GB has been performed on the TAURUS cluster in 245 seconds (approximately 4 minutes) using 391 cores with 10 GB main memory each. Such a rapid segmentation process facilitates a thorough validation procedure and extensive parameter testing concerning the intrinsic model parameters of the RF and CRF.

REFERENCES

Breiman, L., 2001. Random forests. Machine learning, 45(1), pp. 5–32.

Boykov, Y., Kolmogorov, V., 2004. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 26(9), pp. 1124–1137.

Herold, H. 2013. Advances and Recent Challenges in Cartographic Pattern Recognition. In: Proceedings of the 26th ICA International Cartographic Conference, 2013.

Lafferty, J., McCallum, A., Pereira, F. (2001). Conditional random fields: Probabilistic models for segmenting and labeling sequence data. Proc. 18th International Conference on Machine Learning. Morgan Kaufmann. pp. 282–289.

Schemala, D., 2016. Semantische Segmentierung historischer topographischer Karten. Diploma thesis, TU Dresden, 57 p.