Abstract DH2017
Machine Vision algorithms on cadaster plans

Sofia Ares Oliveira, Frederic Kaplan, Isabella di Lenardo {sofia.oliveiraares, frederic.kaplan, isabella.dilenardo}@epfl.ch

Digital Humanities Laboratory, Ecole Polytechnique Fédérale de Lausanne

March 27, 2017

1 Introduction

Cadaster plans are cornerstones for reconstructing dense representations of the history of the city [1]. They provide information about the city urban shape, enabling to reconstruct footprints of most important urban components (buildings, streets, canals, bridges) as well as information about the urban population and city functions (census information, property, rent prices, etc.) [2]. Cadasters plans are usually the results of coordinated campaigns with standardised methods of measurement and representation. This means that large sets of documents follow the same representation conventions. This regularity opens the possibly of efficient automated process for analysing them and possibly transforming the information they contain in georeferenced databases that can be used as part of historical geographical information system [3].
However, as some of these handwritten documents are more than 200 years old, the establishment of processing pipeline for interpreting them remains extremely challenging. This may explain why, to our knowledge, no such system exists in the literature. This article reports our effort in this domain, presenting the first implementation of a fully automated process capable of segmenting and interpreting Napoleonic Cadaster Maps of the Veneto Region dating from the beginning of the 19^{th} century. Our system extracts the geometry of each of the drawn parcels, classifies, reads and interprets the handwritten labels. We believe the general principle of technologies used in the process could be adapted to other cadastral funds, but this has not been tested in the present study.

2 Methodology

Literature on map processing includes works on many different types of maps, from roads to topographic maps, including hydrographic and cadastral maps. Most studies focus on particular problems and features and thus develop techniques that are highly map specific [4].
Our work addresses the particular case of the Napoleonic cadaster of Venice dated 1808, but aims at developing a method highly adaptable to other cadasters with little extra effort.
We propose a system that segments the cadastral map, identifies and extract segmented objects such as parcels and identifiers and recognises the extracted hand-written digits. A demo code with examples of the results can be found at https://github.com/dhlab-epfl/cadasters. The method is

summarized in Fig. 1.

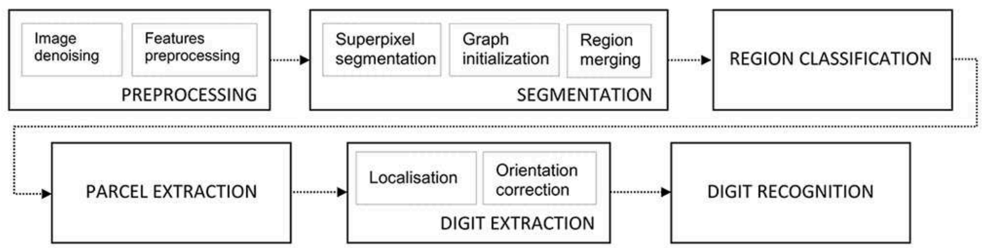


Figure 1: Overview of the system

2.1 Preprocessing

Usually, the processed images are ancient documents that have been digitised. To deal with the natural ageing of paper and eventual spots on the map without losing details, we use a non-local means denoising method [5] to smooth the image.

2.2 Segmentation

We address the task of extracting the desired information from the document as a segmentation problem, which is a recurrent problem in image processing. A graph-based segmentation approach is adopted, which models the image as a weighted undirected graph. This allows to process the pixels or regions in the spatial domain of the image but also to use higher level information such as connections, similarities and dependencies between the elements.
Because a group of pixels sharing some similarities are more perceptually meaningful than a simple pixel, we use SLIC method [6] to create superpixels. Superpixels are clusters of pixels that share similarities and spatial proximity and have the advantage of reducing the complexity of image processing tasks.
A graph is a mathematical structure composed of vertices and edges, representing a system of connections or interrelations among a set of objects. It is widely used to model relations, to study information systems or to organise data. In our case, the graph representing the image is initialized with superpixels as vertices. Its edges connect neighbouring vertices (superpixels) and each edge has a weight which is a measure of the dissimilarity between neighbouring elements. The distance (or dissimilarity) metric is based on color and edge/ridge features.
The oversegmentation of the image resulting from superpixel generation is then reduced by grouping superpixels into homogeneous regions and merging the corresponding graph vertices. Our approach uses global homogeneity, meaning that the method minimize intragroup dissimilarity and maximize intergroup dissimilarity. The 'dispersion' of edge weights (i.e standard deviation within a region) allows to spot high weighted edges within a group and thus disconnect dissimilar vertices (i.e remove their edge) to end up with independent homogeneous regions.

2.3 Region classification

The merged regions are classified into 3 classes : text, contour/delimitations and background (smooth textures such as parcels or streets) using a SVM classifier. The training data is composed of manually annotated samples of maps coming from the Napoleonic cadaster of Venice.

2.4 Parcel extraction

Classification results allow to determine possible parcels candidates and flood fill algorithm is applied, using a ridge detector to indicate boundaries. The chosen ridge detector was originally developed as a vessel enhancement filter [7] and looks for multiscale second order local structures of the image that can be regarded as tubular. The obtained measure indicates how similar the structure is to a tube, and so it is able to detect ridges. Starting from one point in the regions labelled as background (seed point), the flood fill algorithm floods each flat' zone, i.e parcels, streets, etc. and stops at the boundaries (output of ridge detector).
Each parcel of the image is extracted as a polygonal shape and the polygon's corner points are stored in GeoJSON format. If the image file is georeferenced and contains geographical information (GTIFF file for instance), polygons are exported according to the spatial reference system provided. This allows a fast an easy integration of the shapes into a geographic information system (GIS) and thus geographic information on the parcels can easily be collected.

2.5 Digit extraction

Parcel's identifier is usually contained within the parcel. This observation and extracted polygons' information can be used to correct misclassified text regions and improve identifiers extraction.
Elements labelled as text regions are localised, delimited by bounding boxes and grouped so that neighbouring characters are extracted together. Again, information from polygons is used to determine whether neighbouring digits belong to the same identifier or not (i.e whether neighbouring digits are located in the same parcel/polygon). Boxes that do not correspond to identifiers or digits are removed according to some criteria. Finally, the boxes containing parcels' identifiers are extracted.
Since the digit recognition step requires horizontally oriented digits to output accurate prediction, the identifiers' boxes are rotated. Principal analysis component is applied on the binary image of the extracted numbers to determine the angle of the rotation.

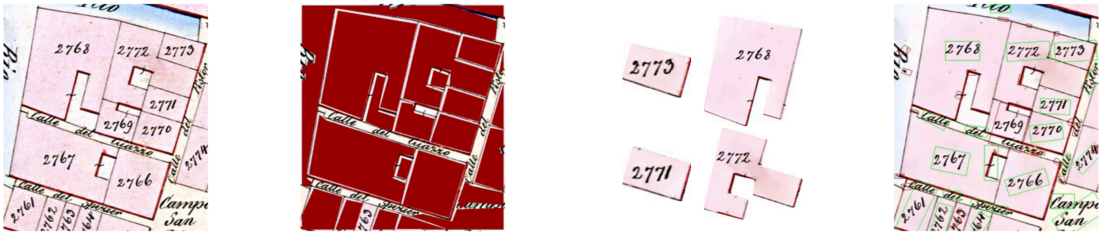
2.6 Digit recognition

The horizontally oriented numbers are separated into digits that are processed individually. A good digit segmentation is primordial since connected or overlapping digits lead to incorrect recognition. A Convolutional Neural Network (CNN) with two convolutional layers, two fully connected layer and a final softmax layer for multiclass classification is used to predict the identifiers. The CNN is trained on a mixed dataset composed of MNIST dataset [8] and digit samples from *Sommarioni* register and has a performance of 99.1%. When predicting the numbers, the network outputs the inferred number with a confidence level, which indicates the reliability of the result.

3 Results

The proposed approach shows promising results in parcel extraction and identifiers recognition. We performed the first 'proof-of-concept' evaluations on manually labelled data taken from different cadaster samples. The total number of annotated object are shown in Table 1.

Most parcels and identifiers were correctly extracted (Table 2 & 3), which comforted us on the feasibility of their automatic extraction. The precision can still be increased for example by using feedback from digit recognition results, i.e the prediction and its confidence level would allow to discard regions where no reliable identifier has been recognised.



(a) (b) (c) (d)

Figure 2: Sample of results : (a) original image, (b) polygon approximation of parcels, (c) extracted parcels and (d) identifiers localization

Parcels with labels	810
All parcels (with and without labels)	1185
Parcels' numbers	736

Table 1: Count of ground-truth objects

Concerning the digit recognition, only around 10% of the identifiers had their digits correctly recognised. Since the models used have shown good performance on nicely detached digits, the fault is not on the recognition algorithm itself but rather on the digit segmentation procedure. The current segmentation is the main hindrance to an efficient digit recognition, thus, further work should focus on a better number processing algorithm. Another alternative is to avoid the segmentation problem and use a recurrent neural network such as LTSM to process the number as a sequence.

4 Perspectives

Our work shows promising results to ease and accelerate cadaster processing, especially with its efficient parcel segmentation and digit identification. Moreover, the export of parcel's geometry into GeoJSON format opens up further perspectives to efficiently georeference ancient maps. The system can be extended and integrated into a user interface to take better advantage from the results, for example by allowing the user to correct or add information about parcels and identifiers.

The proposed method makes a bridge between two data types that were so far separated: the raster object and the vector object. Currently, web-mapping tools consider vector objects as separate layers on the raster maps and each object needs to be manually redesigned. The automatic vectorization process enables to perform the visualisation and annotation processes directly on the cartographic source without the necessity prerequisite of complex geomatics skills. It should greatly facilitate the large scale exploitation of such kind of documents.

	Labelled parcels			All parcels		
IoU	> 0.6	> 0.7	> 0.8	> 0.6	> 0.7	> 0.8
Recall	0.77	0.76	0.72	0.72	0.69	0.60
Precision	0.55	0.54	0.51	0.75	0.71	0.62
Ground-truth	810			1185		
Total extracted	1144					

Table 2: Results of parcel extraction with different Intersection over Union (IoU) thresholds

Inter	> 0.5	> 0.7	> 0.9
Recall	0.90	0.87	0.81
Precision	0.58	0.55	0.51
Ground-truth	736		
Total localized	1152		

Table 3: Results of parcels' number localization with different Intersection (overlapping percentage) thresholds.

	Correct number	Partial number			
		4 digits	3 digits	2 digits	1 digit
MNIST	58 (.09)	17 (.03)	105 (.16)	94 (.14)	165 (0.25)
MNIST-*Sommarioni*	66 (.10)	20 (.03)	90 (.14)	103 (.16)	163 (.26)
Total localized	637				

Table 4: Results of parcels' number recognition

References

[1] I. di Lenardo and F. Kaplan, "Venice time machine: Recreating the density of the past," in *Digital Humanities 2015*, no. EPFL-CONF-214895, 2015.

[2] H. Noizet, B. Bove, and L. Costa, "Paris de parcelles en pixels," 2013.

[3] I. N. Gregory, K. K. Kemp, and R. Mostern, "Geographical information and historical research: Current progress and future directions," *History and Computing*, vol. 13, no. 1, pp. 7–23, 2001.

[4] Y.-Y. Chiang, S. Leyk, and C. A. Knoblock, "A survey of digital map processing techniques," *ACM Computing Surveys (CSUR)*, vol. 47, no. 1, p. 1, 2014.

[5] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 2, pp. 60–65, IEEE, 2005.

[6] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, "Slic superpixels compared to state-of-the-art superpixel methods," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 34, no. 11, pp. 2274–2282, 2012.

[7] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," in *Medical Image Computing and Computer-Assisted Interventation—MICCAI'98*, pp. 130–137, Springer, 1998.

[8] Y. LeCun, C. Cortes, and C. J. Burges, "The mnist database of handwritten digits," 1998.