ISPRS Journal of Photogrammetry and Remote Sensing 95 (2014) 93–108

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

# Tensor-Cuts: A simultaneous multi-type feature extractor and classifier and its application to road extraction from satellite images

Charalambos Poullis [1]

*Immersive and Creative Technologies Lab, Cyprus University of Technology, Cyprus*[2]

ARTICLE INFO

*Article history:*
Received 24 September 2013
Received in revised form 5 June 2014
Accepted 6 June 2014

*Keywords:*
Feature extraction
Feature classification
Graph-Cuts
Road extraction
Tensor

ABSTRACT

Many different algorithms have been proposed for the extraction of features with a range of applications. In this work, we present Tensor-Cuts: a novel framework for feature extraction and classification from images which results in the simultaneous extraction and classification of multiple feature types (surfaces, curves and joints). The proposed framework combines the strengths of tensor encoding, feature extraction using Gabor Jets, global optimization using Graph-Cuts, and is unsupervised and requires no thresholds. We present the application of the proposed framework in the context of road extraction from satellite images, since its characteristics makes it an ideal candidate for use in remote sensing applications where the input data varies widely. We have extensively tested the proposed framework and present the results of its application to road extraction from satellite images.



## 1. Introduction

For many years, there has been a plethora of research work in the area of automatic extraction of road networks from satellite images. Many different algorithms have been proposed so far each with its strengths and weaknesses. Although there have been considerable progress in this area there still exists a rather large gap between the current state-of-the-art and the desired goal. To make matters worse each proposed algorithm excels when used with a particular type of data and fails with another. Quite often, the primary reason for this is due to the fact that existing segmentation frameworks used in the processing pipeline are only able to deal with particular types of imagery, or even worse require an abundance of thresholds which have to be tuned for different types of data/images. Hence, it is no surprise that the segmentation is often identified as the *weakest link* in automatic road extraction pipelines since it dictates the accuracy and quality of the final results.

Motivated by the limitation of our earlier work in Poullis and You (2010) to only classify curve pixels, we present a novel framework for the simultaneous segmentation and classification of multiple image features, called Tensor-Cuts. The proposed framework contains *no thresholds* since *all* information is encoded as tensors, and the efficient and global optimization technique Graph-Cuts is used to refine this information. Our technical contributions are:

- A novel segmentation framework which relies on tensorial representation, Gabor filters and global optimization using Graph-Cuts; named Tensor-Cuts.
- The proposed segmentation has *no threshold* since *all* information is encoded as tensors and refined using Graph-Cuts.
- An incremental improvement on the road center-point extraction procedure for the automatic extraction of center-points.

In contrast to our previous work (Poullis and You, 2010), the proposed segmentation encodes labels as tensors which can *simultaneously* capture information about three geometric types: surface, curve and junction. This eliminates the need for a pre-classification step (to separate curve pixels) prior to the segmentation as was the case in Poullis and You (2010) since both, the initial and final labeling have the form of tensors. Moreover, the energy function penalizes for dissimilarities between *all* pixels/tensors (rather than differences between orientations of *only* the pixels classified as curves) resulting in better defined segmentation. Thus, the user interaction previously required to mark foreground and background areas in the image prior to the processing is no longer required making the proposed framework automatic. Furthermore, the newly introduced energy label term in the Graph-Cut optimization ensures that a *minimum* number of labels are used in the new labeling which produces smoother results. Finally, the

[1] More information: http://www.poullis.org.
[2] http://www.theictlab.org.
*E-mail address:* charalambos@poullis.org

http://dx.doi.org/10.1016/j.isprsjprs.2014.06.006
0924-2716/

extraction of roads is achieved using an improved and more efficient variant of the procedure described in Poullis and You (2010) which exploits prior-knowledge about the characteristics and the function of roads to efficiently apply a set of Gaussian kernels to extract road *entry points* in the image. Once the entry points have been identified a variety of road tracking algorithms can then be employed for extracting the complete network; in our case tracking using template matching.

The paper is organized as follows: Section 2 presents a brief overview of the state-of-the-art in the area and Section 3 a system overview of the proposed method. Section 4 explains the intrinsic details of how the encoding of the information into their tensorial representation is performed. This includes the application of Gabor filters (Section 4.1.1), and the conversion of the response images to tensors (Section 4.1), the creation of the tensor labels (Section 4.2). Section 5 presents the application of the global optimization Graph-Cuts using tensors and an explanation of the energy functions used. Section 6 presents the road extraction pipeline which incorporates the segmentation framework. Finally, Section 7 shows the results of the proposed technique.

## 2. Related work

There has been considerable work in the area of automatic road extraction. Below we present an overview of the state-of-the-art in the area.

Doucette et al. (2001) present a self-organizing road map algorithm. Their approach proceeds by performing spatial cluster analysis as a mid-level processing technique. This enables them to improve tolerance to road clutter in high-resolution images, and to minimize the effect on road extraction of common classification errors. Their approach is designed in consideration of the emerging trend towards high-resolution multispectral sensors. Indeed, their preliminary results demonstrate robust road extraction ability due to the non-local approach, when presented with noisy input.

Hinz and Baumgartner (2003), propose a system which integrates detailed knowledge about roads and their context using explicitly formulated scale-dependent models. The knowledge about how and when certain parts of the road and context model are optimally exploited is expressed by an extraction strategy.

Bacher and Mayer (2005), proposed an automatic road extraction technique from satellite images of rural and suburban areas. The first step is the extraction of Steger lines in all spectral channels which are then used as cues for roads to generate training areas for a subsequent automatic supervised classification. The resulting classification is finally used as an additional source for the extraction of road candidates.

Mena and Malpica (2005) propose a system which includes four different modules: data pre-processing; binary segmentation based on three levels of texture statistical evaluation; automatic vectorization by means of skeletal extraction; and finally a module for system evaluation. The proposed system is quite efficient in producing the results.

On a different note, Lacoste et al. (2005) propose an approach which models the target line network by an object process, where the objects correspond to interacting line segments. They design the prior model so that they will exploit as fully as possible the topological properties of the network under consideration, while the radiometric properties of the network are modeled using a data term based on statistical tests.

In a different approach Grote and Heipke (2008), proposed a region-based approach on high resolution aerial images working from small local regions to roads as groups of road parts. The first step is to segment the image using the normalized cuts algorithm and group small segments to form larger segments; from these grouped segments road parts are then extracted.

In Das et al. (2011), the authors propose a multi-stage automated system for extracting road networks from high-resolution satellite images which produces impressive results. During the first stage, information about edges is extracted using a proposed process called "dominant singular measure" and information about regions is extracted using a probabilistic SVM classifier. This information is then integrated together using a constraint satisfaction neural network. A set of post-processing steps is then applied at a second stage, in order to improve the accuracy the extracted roads by removing false positives as well as to recover missing small areas due to false negatives. The system does not vectorize the road network but rather provides a binary classification of each pixel into road or non-road.

By employing neural networks with millions of trainable weights, Mnih and Hinton (2010) proposed the detection of roads which looks at a much larger context than was used in previous attempts at learning the task. The network is trained on massive amounts of data using a consumer GPU. The results confirm that their method works reliably on challenging urban datasets that are an order of magnitude larger than what was used to evaluate previous approaches.

Using stereoscopic satellite aerial images, Poz et al. (2012) proposes a semiautomatic method for 3-D road extraction in rural areas. They employ a strategy based on the dynamic programming algorithm which provides a solution to the road extraction problem in the object space. In order to find road centerlines, the extraction process begins by first measuring a few seed points in one image of the stereoscopic pair and then transforming these into the object-space reference system. Experimental results show that the proposed method is efficient and provides relatively accurate road centerlines.

A higher-order CRF model is presented in Wegner et al. (2013) for road network extraction from dense urban scenes. A CRF formulation is proposed for road labeling, where the prior is represented by higher-order cliques which connect sets of superpixels along straight line segments. Although the parameters are manually tuned for the clique sampling the results seem very promising.

Although a plethora of techniques have been proposed for the automatic or semi-automatic road extraction, the gap between the state-of-the-art and the desired goal, of automatic road extraction from satellite images, still remains wide. In this work we introduce a framework which is fully automatic and does not require interaction with the user to mark road and/or non-road pixels nor to exactly separate various objects from the image.

## 3. System overview

Fig. 1 shows an overview of the proposed system consisting of two phases:

1. *Tensor encoding:* The original image is filtered with a set of Gabor Jets and the responses are encoded into tensors. Similarly, a Gaussian hemisphere's geometric properties are encoded as tensors as described in Section 4.
2. *Global optimization:* The encoded tensors are optimized using the efficient, global optimization technique, Graph-Cuts, which is explained in Section 5.
3. *Road extraction:* Road center-point candidates are extracted using a set of single and bi-modal Gaussian-based kernels. Based on the road center-points the roads are then extracted via tracking.