Coastal Engineering 105 (2015) 1–12

Contents lists available at ScienceDirect

Coastal Engineering

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

An automated method for semantic classification of regions in coastal images

B.M. Hoonhout [a,b,*], M. Radermacher [a], F. Baart [a,c], L.J.P. van der Maaten [d]

[a] Department of Hydraulic Engineering, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1, 2628CN Delft, Netherlands
[b] Unit of Hydraulic Engineering, Deltares, Boussinesqweg 1, 2629CN Delft, Netherlands
[c] Deltares Software Centre, Deltares, Boussinesqweg 1, 2629CN Delft, Netherlands
[d] Department of Intelligent Systems, Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Mekelweg 4, 2628CD Delft, Netherlands

ARTICLE INFO

Article history:
Received 29 January 2015
Received in revised form 12 June 2015
Accepted 4 July 2015
Available online 2 September 2015

Keywords:
Coastal
Images
Semantic classification
Support vector machine
Features

ABSTRACT

Large, long-term coastal imagery datasets are nowadays a low-cost source of information for various coastal research disciplines. However, the applicability of many existing algorithms for coastal image analysis is limited for these large datasets due to a lack of automation and robustness. Therefore manual quality control and site- and time-dependent calibration are often required. In this paper we present a fully automated algorithm that classifies each pixel in an image given a pre-defined set of classes. The necessary robustness is obtained by distinguishing one class of pixels from another based on more than a thousand pixel features and relations between neighboring pixels rather than a handful of color intensities.

Using a manually annotated dataset of 192 coastal images, a SSVM is trained and tested to distinguish between the classes *water*, *sand*, *vegetation*, *sky* and *object*. The resulting model correctly classifies 93.0% of all pixels in a previously unseen image. Two case studies are used to show how the algorithm extracts beach widths and water lines from a coastal camera station.

Both the annotated dataset and the software developed to perform the model training and prediction are provided as free and open-source data sources.

1. Introduction

Coastal imagery nowadays is a valuable and low-cost source of information for coastal research in a variety of disciplines. Characteristics such as beach width, water line dynamics, wave breaking and runup, vegetation cover, aeolian dynamics and beach usage are visible to the naked eye from a simple coastal image (e.g. Fig. 1, example, A). Further analysis of the acquired images can provide us with derived information like bathymetries, flow patterns and sediment transport trajectories. Coastal image analysis is not restricted to ordinary visible light imagery, but can be applied to (near-)infrared, multi- or hyperspectral imagery and video as well, increasing the number of coastal features that can be distinguished.

Since investments to install a coastal camera station and corresponding data storage are low compared to most other monitoring alternatives, the amount of coastal camera stations worldwide is increasing steadily. With the increasing amount of coastal imagery data, an increasing number of coastal image analysis algorithms is being developed with a variety of applications. Pioneering work on swash runup estimates from coastal images was done by Holland and Holman (1991). The extraction of runup lines inspired many authors to map intertidal bathymetries from series of runup lines obtained from a series of coastal images (e.g. Plant and Holman, 1997; Aarninkhof et al., 2003; Quartel et al., 2006; Plant et al., 2007; Uunk et al., 2010; Osorio et al., 2012). Many shoreline extraction algorithms are available, including those that use multispectral images for increased precision (e.g. Sekovski et al., 2014). Subsequently, coastal images were used to estimate surfzone currents (e.g. Holland et al., 2001; Chickadel et al., 2003) and later subtidal bathymetries (e.g. Aarninkhof et al., 2005; van Dongeren et al., 2008; Holman et al., 2013). The global presence of coastal camera stations makes it possible to monitor long-term coastal behavior (Smit et al., 2007) and sparked several applications for coastal zone managers, like estimating coastal state indicators (Davidson et al., 2007), deriving beach usage statistics (Jimenez et al., 2007; Guillén et al., 2008), tracking sandbar positions (Lippmann and Holman, 1989; van Enckevort and Ruessink, 2001; Price and Ruessink, 2011) and rip current detection systems (Bogle et al., 2000; Gallop et al., 2009).

The size of the coastal imagery archive grows rapidly. New camera stations are deployed every year, adding to the diversity of the total dataset. The data intake per station is increasing, which makes the total dataset harder to analyze. The applicability and performance of these algorithms on the large coastal imagery datasets that are

* Corresponding author at: Department of Hydraulic Engineering, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1, 2628CN Delft, Netherlands.
E-mail addresses: bas.hoonhout@deltares.nl, b.m.hoonhout@tudelft.nl (B.M. Hoonhout).

http://dx.doi.org/10.1016/j.coastaleng.2015.07.010
0378-3839/

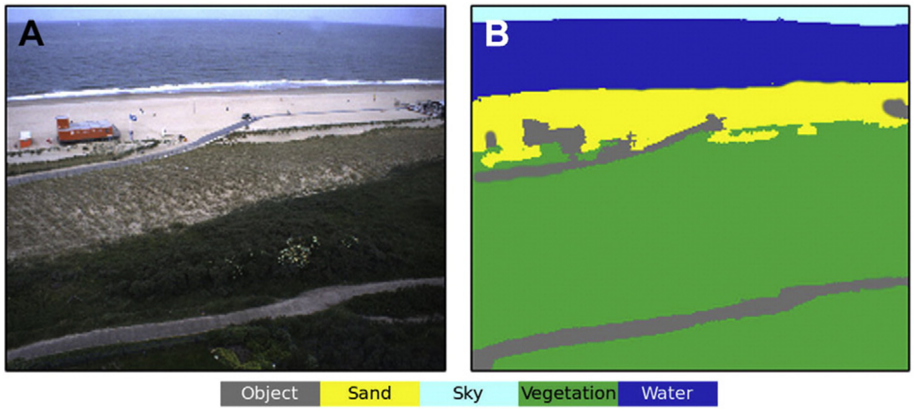

Fig. 1. Example of a coastal image taken on July 1st, 2013 in Kijkduin, The Netherlands (A) and the corresponding manual annotation (B).

presently available depend on two characteristics in particular: automation and distinctiveness of relevant pixels. Many algorithms need some kind of manual pre-selection of relevant pixels (the region of interest) or manual quality control, which often hampers analysis of large, long-term coastal imagery datasets (e.g. Holland and Holman, 1991; Aarninkhof et al., 2003; Jimenez et al., 2007). Algorithms that do not rely on a manual quality control need to distinguish between classes of pixels based on a limited number of intrinsic pixel features, usually the color channels RGB or HSV (e.g. Aarninkhof et al., 2003; Quartel et al., 2006). Consequently, feature characteristics for different classes are very likely to overlap and automated classification of large, long-term coastal imagery datasets becomes unreliable if not unfeasible without site- and time-dependent calibration, limiting the applicability of the algorithm.

To the best of our knowledge, this paper presents the first fully automatic, high-quality algorithm for the supervised segmentation of coastal images into image regions containing major classes such as *water*, *sand*, *vegetation*, *sky*, and *objects* (Fig. 1, B). In contrast to prior work, this algorithm does not rely on a few color features with limited discriminability between classes, but aggregates over more than a thousand features that contain information on color, texture, and visual appearance. In addition, the algorithm uses a machine learning framework that allows us to leverage thousands of features for high-accuracy classification and enables the use of structured learning where relations between neighboring pixels are taken into account. The algorithm is not tailored to any specific classification task, but is merely a general classification framework that can be applied on large, long-term coastal imagery datasets or on parts of individual images. Nonetheless, it produces substantially better results than obtained by tailored algorithms that rely on color features alone and omit the use of structured learning. In addition, we present a manually annotated dataset of coastal images that can be used for training and testing of machine-learning based systems such as ours, and we present an open-source Python toolbox for performing coastal image analysis tasks.

2. Methodology

Automated classification of regions in coastal images is done using a classification model. Classifying image regions into various meaningful classes based on a set of properties (features) shows similarities to regression models (e.g. linear regression). Regression models are used to predict the value of a target parameter based on some input samples. In principle, classification models are regression models, which use a threshold value for the target parameter to distinguish between a set of discrete classes.

Supervised classification models (as opposed to unsupervised models, which are not treated here) require training. During training the optimal threshold values are determined based on an annotated dataset. Optimization is done by minimizing a cost function. The definition of this cost function is the main factor that distinguishes between various model types. For example, a linear regression model usually minimizes the mean squared error of the predicted target parameter over all training samples. A network of regression models, like an artificial neural network, is occasionally used for classification purposes (e.g. Kingston, 2003; Verhaeghe et al., 2008). In this study a method closely related to a LR is used (e.g. Vincent, 1986; Dusek and Seim, 2013). A LR is, although the name suggests regression, a classification method that optimizes the logistic loss function.

The workflow adopted in this study consists of four steps: 1. a manually annotated dataset of coastal images is oversegmented into *superpixels*; 2. for all images in the dataset an extensive set of features is extracted; 3. a suitable classification model is trained using the manually annotated data; and 4. the trained model is used to automatically classify future images. The workflow is visually presented in Fig. 2. In this section these four steps are described. In the next section the performance as well as a first application of the algorithm is presented.

2.1. Dataset

The ArgusNL dataset with manually annotated coastal images consists of 192 images obtained from 4 coastal camera stations located along the Dutch coast (Egmond, Jan van Speijk, Kijkduin and Sand Motor), each containing 5 to 8 cameras.

Each camera that is part of the coastal camera stations used in this study takes a snapshot image twice an hour. Apart from snapshot images, also 10 min mean, variance, min and max images are stored, but these are not used for classification. Also, any images that were obscured and thus did not contain any valuable data are discarded. Images can be obscured either because the image was taken before sunrise or after sunset or because of the presence of rain, fog, sun glare in the water or dirt on the camera lens.

From all suitable snapshot images taken during the summer of 2013 by these cameras 192 images are randomly selected. The images are automatically oversegmented (see Section 2.2) and one of the following classes is assigned manually to each 1. sky; 2. water (sea); 3. water (pool); 4. sand (beach); 5. sand (dune); 6. vegetation; 7. object (sea); 8. object (beach); and 9. object (dune). In this study the classes are aggregated to the most relevant ones, being: 1. sky; 2. water; 3. sand; 4. vegetation; and 5. object.

The ArgusNL dataset, including the images, oversegmentation and annotation, was published by Hoonhout and Radermacher (2014).