ISPRS Journal of Photogrammetry and Remote Sensing 108 (2015) 94–112

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Orientation-selective building detection in aerial images

Andrea Manno-Kovacs *, Tamas Sziranyi

Distributed Events Analysis Research Laboratory, Institute for Computer Science and Control, MTA SZTAKI, Hungarian Academy of Sciences, H-1111, Kende u. 13-17, Budapest, Hungary

ARTICLE INFO

Article history:
Received 3 March 2015
Received in revised form 23 June 2015
Accepted 23 June 2015

Keywords:
Orientation selectivity
Modified Harris for edges and corners
Building detection
Active contour

ABSTRACT

This paper introduces a novel aerial building detection method based on region orientation as a new feature, which is used in various steps throughout the presented framework. As building objects are expected to be connected with each other on a regional level, exploiting the main orientation obtained from the local gradient analysis provides further information for detection purposes. The orientation information is applied for an improved edge map design, which is integrated with classical features like shadow and color. Moreover, an orthogonality check is introduced for finding building candidates, and their final shapes defined by the Chan–Vese active contour algorithm are refined based on the orientation information, resulting in smooth and accurate building outlines. The proposed framework is evaluated on multiple data sets, including aerial and high resolution optical satellite images, and compared to six state-of-the-art methods in both object and pixel level evaluation, proving the algorithm's efficiency.

1. Introduction

Automatic building detection is currently a relevant topic in aerial image analysis, as it can be an efficient tool for accelerating many applications, like urban development analysis and map updating, also providing great support in crisis and disaster management, and in aiding municipalities in long-term residential area planning. Large, continuously changing areas have to be monitored periodically, requiring a huge effort if performed manually. Therefore, there is a high interest for automatic processes to facilitate such analysis.

A wide range of publications is available in remote sensing for building detection for sparsely located building objects, often based on shape estimation or contour outlining. Earlier works like Huertas and Nevatia (1988) introduced a technique for detecting buildings with rectangular components and shadow information. Line based segmentation techniques, like Lin and Nevatia (1998) were based on the extraction of line segments, processed with various methods. Following this principle, Unsalan and Boyer (2005) proposed an extension, where the street network was extracted from the segmented images and houses were detected based on graph theoretical algorithms. In the same manner, Sirmacek and Unsalan (2008) – denoted by *BoxFit* in the experiments – fused shadow and invariant color features with edge information in a two-step process. First, a building candidate was defined based on color and shadow features, then a rectangle was fitted using a Canny edge map. This sequential method was very sensitive to the deficiencies of both steps: inappropriate shadow and color information causing false candidates and inexact edge maps causing inaccurate detections. As the proposed method uses similar information sources, we can highlight the impacts of our contributions by direct comparisons during the evaluation.

Following the region-based trend, Song et al. (2006) introduced a segment-merge technique (*SM*), which considered building detection as a region level task and assumed buildings to be homogeneous areas (considering either color or texture). First, a building model prior was constructed with texture and shape features from a training building set. After selecting building-like regions, shape and size constraints were used to merge such regions into building candidates, followed by shadow and geometrical rules to finalize candidates. However, the basic assumptions influenced the success of the whole approach: when buildings could not be distinguished from the background by using color and texture features, further steps would also fail. Moreover, they assumed simple building models, so complex shapes could not be reconstructed. The orientation of a candidate building region seed was introduced as a useful feature, defining potential rectangle orientations.

A point process based technique was introduced in Ortner et al. (2008), which used stochastic geometry based on the superposition of segment and rectangle processes. The work of Katartzis

* Corresponding author. Tel.: +36 12796106.
E-mail addresses: andrea.manno-kovacs@sztaki.mta.hu (A. Manno-Kovacs), sziranyi@sztaki.mta.hu (T. Sziranyi).

http://dx.doi.org/10.1016/j.isprsjprs.2015.06.007
0924-2716/

and Sahli (2008) is based on a stochastic image interpretation model and applies a Markov random field model to describe the dependencies between the available hypotheses.

Latest publications can be grouped into hierarchical and graph model based approaches: A hierarchical approach was introduced in Benedek et al. (2012), using a multitemporal Marked Point Process (MPP) model combined with a bi-layer Multiple Birth and Death (*bMBD*) optimization for rectangular building detection. Object-level features (exploiting low level features) were integrated into a configuration function, which was then evaluated by a bMBD stochastic optimization process. The result of the process was a group of rectangles, representing detected buildings. Although the hierarchical approach of the method was able to handle diverse objects, it was limited by the applied features (gradient, color, shadow), therefore it had problems when detecting objects with weak features, like non-red roofs. Moreover, the applied strictly rectangular templates prevented the accurate detection of complex shapes.

A graph model based algorithm for polygonal building shape detection was developed in Izadi and Saeedi (2012), employing lines, line intersections, and their relations. Ok et al. (2013) introduced the *GrabCut* partitioning algorithm for building extraction. The algorithm first investigated the shadow evidence to select potential building regions by applying a fuzzy landscape generation approach to model the directional spatial relationship between buildings and their shadows (which motivated our orientation selective fusion step in the proposed method). Then a pruning process was developed to eliminate non-building objects. Finally, *GrabCut* partitioning detected the building regions. However, a drawback of the algorithm was its sensitivity to the shadow extraction step, therefore it had problems when detecting buildings without shadow or having only fragmented shadow parts.

The method presented here is based on the fact that feature point detectors can be applied efficiently for man-made object detection, also indicated in Martinez-Fonte et al. (2005), where Harris and SUSAN detectors, published in Harris and Stephens (1988) and Smith and Brady (1997), were validated for distinguishing man-made versus natural structures. The principle was followed also in Peng et al. (2005) using the Tomasi and Kanade corner detector and in Sirmacek and Unsalan (2009), introducing a graph construction approach (denoted by *SIFT-graph*) for urban area and building detection using SIFT keypoints proposed in Lowe (2004). The method applied a light and a dark template to represent buildings. First, SIFT feature points were extracted from the image, followed by graph based techniques to detect urban areas. The given templates helped to divide the point set into separate building subsets and to define the locations of the different objects without any shape estimation. However, in many cases, not all building objects could be represented by only two templates, moreover, the given features were not always enough to distinguish the buildings from the background.

Cui et al. (2008) and Cote and Saeedi (2013) introduced a method using Harris corner points; Sirmacek and Unsalan (2011) tested directional Gabor filter based feature points, Harris corner points and the FAST points of Rosten et al. (2010) for extracting different local feature vectors (*LFV*), estimating a joint probability density function for urban area detection, assuming that around such points there is a high probability for urban characteristics. This technique motivated our previous work in Kovacs and Sziranyi (2013) for introducing the Modified Harris for Edges and Corners (MHEC) feature point set for efficient urban area detection.

The main contribution of the present approach is the application of orientation as a novel feature in a direction-selective framework for detecting buildings. Earlier approaches like Sirmacek and Unsalan (2008) and Ok et al. (2013) usually dealt with orientation in terms of the illumination angle. Others applied orientation based techniques for indirect solutions, like Unsalan and Boyer (2004) to identify line support regions. Ortner et al. (2008) introduced the alignment interaction in the MRF energy term. Cui et al. (2012) followed a novel interpretation of orientation information, by using perpendicular building borders to define dominant directions. They were looking for lines in Hough space with orientations indicating potential straight building borders. Similarly, Benedek et al. (2012) extracted a local gradient orientation density function, published in Kumar and Hebert (2003), to find the main orientations characterizing a building and measuring the orthogonality of the candidate area.

However, these approaches concentrated only on one building and its neighborhood to perform the directional analysis, while our proposed method handles orientation as a region level feature and combines it with other features throughout the approach (see Fig. 1):

- The orientation feature is calculated for the whole image and an **improved edge map** is defined, which emphasizes edges only in the main orientations. This improved edge map is fused with color and shadow features, resulting in accurate localization of building candidates.
- An **orthogonality checking** of the possible building blobs is applied to find remaining candidates with limited feature evidence.
- A **directional refinement step** is performed after the Chan–Vese active contour based building detection phase: the shapes of the final blobs are refined by a novel directional operator to efficiently smooth the irregularities of the active contour outline.

2. Proposed framework

In this paper, a novel framework, called Orientation Selective Building Detection (**OSBD**) is proposed, based on the exploitation of multiple features: color, shadow map, and a novel orientation descriptor. Our guiding principle is that, when detecting buildings, orientation is a valuable information, as the alignment of buildings usually depends on the structural properties of their surroundings (e. g. the road network). Therefore the main edges of neighboring buildings should have similar orientations. The main contribution of this paper is to introduce region orientation as a novel feature for building detection and to use this information in a direct manner, unlike previous works e. g. Ortner et al. (2008) where only alignment interaction was calculated between building candidates, avoiding the use of the orientation value itself.

Moreover, we exploit orientation information in multiple steps: the directional descriptor also helps in the verification of the building candidates in the detection step. An orthogonality check is created to validate whether the candidate is a real building or some other image structure. Finally, a directional morphological operator is applied for the remaining building blobs to smooth the final outline, resulting in more accurate detection results.

The main steps of the proposed method with the corresponding subsections are shown in Fig. 1. The two main parts are Feature extraction (Section 2.1) and Feature fusion and contour detection (Section 2.2). As a first step, a feature point set is extracted which is based on the modification of the Harris corner detector (Section 2.1.1), proposed in Kovacs and Sziranyi (2012a). This point set is used as a directional sampling set to compute orientation statistics in Section 2.1.2. Local orientation information is calculated as the main orientation of gradients in the close proximity of the feature points, and extended to the whole image to produce a directional map. Using this map, dominant directions describing the urban regions are defined, helping the construction of a more accurate edge map in Section 2.1.3 by specifying the favorable edge