

BUILDING DETECTION USING DIRECTIONAL SPATIAL CONSTRAINTS

H. Gökhan Akçay, Selim Aksoy

Department of Computer Engineering
Bilkent University
Bilkent, 06800, Ankara, Turkey
{akçay,saksoy}@cs.bilkent.edu.tr

1. INTRODUCTION

Automatic detection of buildings in very high spatial resolution remotely sensed imagery has been an important problem because the detection results can be used in many applications such as change detection, urbanization monitoring, and digital map production. For example, as one of the most salient features of human settlements, precise identification and localization of buildings provide key information sets needed for territorial planning and in any assessment related to human security such as preparedness to natural hazards and to post-disaster evaluation [1]. Furthermore, human settlement analysis for slum and unorganized settlement monitoring can be assisted by automatically extracted building information because slum areas can generally be characterized by a high density of short and small buildings in irregular spatial arrangements [2, 3]. Similarly, buildings can be considered as one of the best indicators for human population estimation.

There is an extensive literature on building detection where both pixel level and object/region level processing have been used. However, most of the previous methods try to solve the problem for specific settings such as images having buildings with the same type of appearance and images where the buildings are isolated and have simple roof structures. With the increase in the spatial details in the images obtained from new generation sensors with meter and sub-meter spatial resolution, the buildings may have very complicated appearances and may have complex structures with very different spectral signatures. Popular edge/line-based and morphology-based approaches also do not often work for complex urban scenes because the contrast among the parts of a roof can be higher than the contrast between the roof and its surroundings (as shown in the examples in Figure 1).

Even though different buildings may appear in significantly different colors and shapes, a common property of such buildings can be the existence of shadows. The relationship between buildings and shadows has actually been exploited in earlier works [4, 5]. More recently, Sirmacek and Unsalan [6] detected buildings with red roofs using color information and verified their existence with the occurrences of shadow-like nearby regions. However, the assumption of red roofs is limiting and there may be other sources of shadows in the image.

This paper proposes a method for detection of buildings with complex shapes and roof structures in very high spatial resolution images by exploiting spectral, structural, and contextual information using a mathematical morphology-based context model and minimum spanning tree-based clustering. First, watershed segmentation is

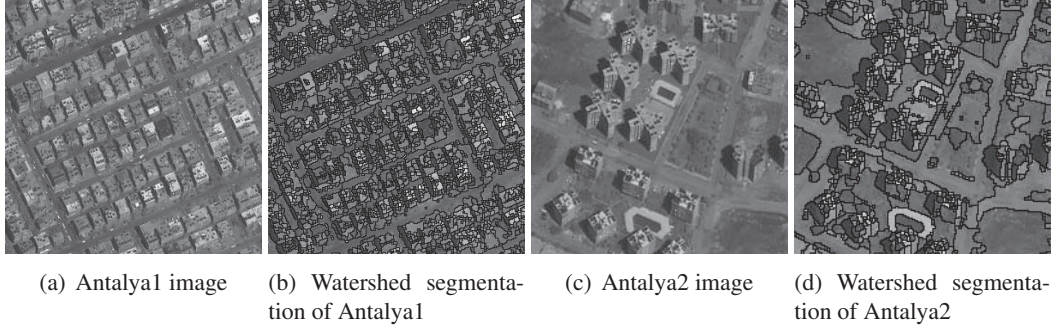


Fig. 1. Examples from an Ikonos panchromatic image of Antalya, Turkey and the corresponding watershed segmentation results. The segmentation boundaries are overlaid as black.

applied to obtain oversegmented regions. Then, shadow regions are detected in this oversegmentation based on their spectral properties. Next, candidate building regions are identified using the directional spatial relationships of all regions with respect to the detected shadow regions along the illumination direction. Finally, the building regions are selected by clustering the oversegmented regions that satisfy the spatial constraints using minimum spanning trees. Experiments are performed using Ikonos images.

2. IMAGE SEGMENTATION AND SHADOW REGION DETECTION

Image segmentation is performed using the classical watershed segmentation algorithm to partition the image into spectrally homogeneous regions. The results contain oversegmented regions because the test areas in this study include buildings with complex roof structures as shown in Figure 1. Other segmentation methods can also be used but similar results are likely to be obtained because of the complex spectral appearance within building regions.

Among all regions, the ones that are likely to belong to shadows are selected using their spectral properties. First, the normalized difference vegetation index (NDVI) is computed using the pan-sharpened image. Then, the regions whose average brightness values are lower than a threshold and average NDVI values are lower than a threshold are denoted as shadow regions. More complicated shadow detection methods can also be used but the aforementioned method performed sufficiently well in the initial experiments.

3. DIRECTIONAL SPATIAL CONSTRAINTS

The candidate building regions are identified by using the shadow regions as directional spatial constraints in a model that we recently proposed for contextual classification and retrieval [7]. Given a reference object B and a direction specified by the angle α , the landscape $\beta_\alpha(B)$ around the reference object along the given direction can be defined as a fuzzy function from the image space \mathcal{I} into $[0, 1]$. The fuzzy membership value $\beta_\alpha(B)(x)$ of an image point $x \in \mathcal{I}$ corresponds to the degree of its satisfaction of the directional spatial relation relative to the reference object B .

In [7], we proposed to compute the fuzzy landscape using the morphological dilation of B ,

$$\beta_\alpha(B)(x) = (B \oplus \nu_{\alpha, \lambda, \tau})(x) \cap B^c, \quad (1)$$

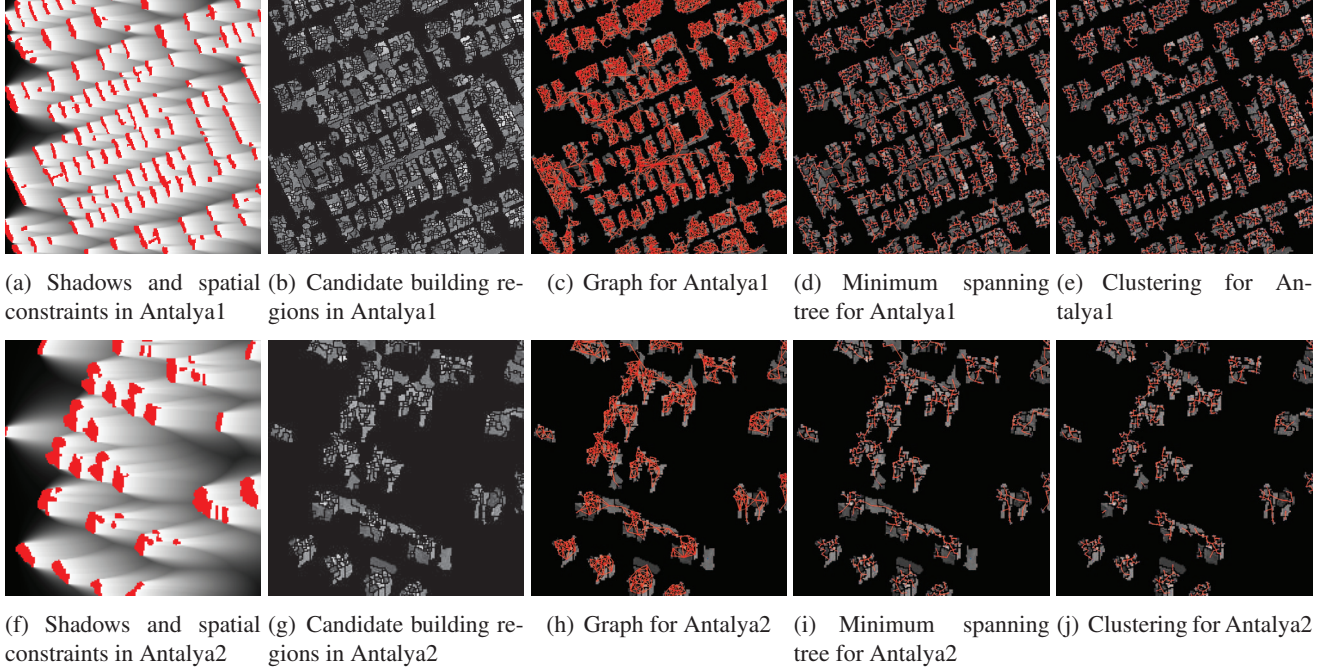


Fig. 2. Examples of shadow regions, directional landscapes, candidate building regions, graph construction, and minimum spanning tree-based clustering.

using the fuzzy structuring element

$$\nu_{\alpha,\lambda,\tau}(x) = g_{\lambda} \left(\frac{2}{\pi} \theta_{\alpha}(x, o) \right) \max \left\{ 0, 1 - \frac{\|\vec{ox}\|}{\tau} \right\} \quad (2)$$

where o is the origin (center) of the structuring element, $\theta_{\alpha}(x, o)$ is the angle measured between the unit vector along the direction α with respect to the horizontal axis and the vector from o to the image point x , $g_{\lambda}(\cdot)$ is a nonlinearly decreasing function with the shape of a Bézier curve, and $\|\vec{ox}\|$ is the Euclidean distance of point x from o . The function g decreases the degree of the relationship as the angle θ increases when the point x departs from α (λ models the extent of the decrease). The second part of (2) decreases the degree of the point's spatial relation to the reference object according to its distance to that object where τ is a threshold corresponding to the distance where a point is no longer visible from the reference object. This definition provides a structuring element that is tunable along both angular and radial dimensions (see [7] for more details).

Given the sun illumination direction, we can find the directional landscapes of the shadow regions along this direction by using (1). The resulting directional landscapes give high responses in areas close to the shadow regions along the sun illumination direction. These areas correspond to the locations where the probability of the presence of buildings is high. Figures 2(a) and 2(f) show the shadow regions and the corresponding landscapes. Consequently, the regions whose average satisfaction degrees are higher than a threshold, average NDVI values are lower than a threshold, and sizes are lower than a threshold are identified as candidate building regions. Figures 2(b) and 2(g) show examples for candidate regions. As can be seen from the figures, most of the regions are correctly identified with a small number of misdetections and several false alarms.

4. GRAPH-THEORETIC BUILDING MODEL

After obtaining the candidate regions, our aim is to identify the regions corresponding to building parts. An important observation is that regions forming a building are densely located whereas regions separating different buildings are found far from their neighbors. The distance between two regions is measured as the distance between their centroids. This seems to be a valid assumption because the regions are obtained from oversegmentation and mostly have compact shapes. Hence, we construct a graph where the graph nodes correspond to the candidate regions' centroids and the edges are created between two neighboring nodes. What we expect is that the nodes representing parts of building regions will form dense subgraph components.

After constructing the graph, the goal is to group the regions into clusters so that each group corresponds to a building or a non-building area. Therefore, we assign a weight to each edge as the spatial distance between the corresponding nodes. Then, to determine the most relevant neighbors of each node, we construct the minimum spanning tree of the graph by using these edge weights. By constructing the tree, a node is connected to its most important and most related neighbors while its relationships with the neighbors that are further away can be ignored.

To cluster the nodes into groups, some edges of the minimum spanning tree should be removed. This is achieved by removing the edges that are longer than a threshold. As a result, the nodes that are spatially close enough remain in the same cluster. Figure 2 shows examples for graph construction and clustering. It can be seen that most of the building regions that cannot be obtained by traditional spectral segmentation methods that cannot incorporate structural and contextual information are correctly extracted by the proposed method.

Once the building regions are detected, they can be used to improve scene analysis [8] and urban area classification [2]. The final version of the paper will include improvements to the minimum spanning tree-based clustering algorithm and performance evaluation with an analysis of the sensitivity of the results with respect to the thresholds.

5. REFERENCES

- [1] M. Pesaresi, A. Gerhardinger, and F. Kayitakire, "A robust built-up area presence index by anisotropic rotation-invariant textural measure," *IEEE JSTARS*, vol. 1, no. 3, pp. 180–192, 2008.
- [2] E. Dogrusoz and S. Aksoy, "Modeling urban structures using graph-based spatial patterns," in *IGARSS*, 2007.
- [3] M. Stasolla and P. Gamba, "Spatial indexes for the extraction of formal and informal human settlements from high-resolution SAR images," *IEEE JSTARS*, vol. 1, no. 2, pp. 98–106, 2008.
- [4] A. Huertas and R. Nevatia, "Detecting buildings in aerial images," *Computer Vision, Graphics, and Image Processing*, vol. 41, no. 2, pp. 131–152, 1988.
- [5] R. B. Irvin and D. M. McKeown Jr, "Methods for exploiting the relationship between buildings and their shadows in aerial imagery," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 19, no. 6, pp. 1564–1575, 1989.
- [6] B. Sirmacek and C. Unsalan, "Building detection from aerial images using invariant color features and shadow information," in *ISCIS*, 2008.
- [7] S. Aksoy and R. G. Cinbis, "Image mining using directional spatial constraints," *IEEE Geoscience and Remote Sensing Letters*, 2010, to appear.
- [8] H. G. Akcay and S. Aksoy, "Automatic detection of geospatial objects using multiple hierarchical segmentations," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 7, pp. 2097–2111, 2008.