Building Detection from Aerial Images using Invariant Color Features and Shadow Information

Beril Sırmaçek and Cem Ünsalan
Computer Vision Research Laboratory
Department of Electrical and Electronics Engineering
Yeditepe University
İstanbul, 34755 TURKEY
e-mail:unsalan@yeditepe.edu.tr

***Abstract*—Robust detection of buildings is an important part of the automated aerial image interpretation problem. Automatic detection of buildings enables creation of maps, detecting changes, and monitoring urbanization. Due to the complexity and uncontrolled appearance of the scene, an intelligent fusion of different methods gives better results. In this study, we present a novel approach for building detection using multiple cues. We benefit from segmentation of aerial images using invariant color features. Besides, we use the edge and shadow information for building detection. We also determine the shape of the building by a novel method.**

I. Introduction

Detecting buildings from aerial images is becoming an important problem in computer vision and remote sensing. One possible reason is that, this detection leads to automatic land mapping applications. However, aerial images have diverse properties that make it hard to develop generic algorithms for building detection. These images also differ in scale (resolution), sensor type, orientation, quality, and ambient lighting conditions. In addition to these difficulties, buildings may have complicated structures and can be occluded by other buildings or trees. Therefore, one has to consider both aspect, structural, and deterministic cues to construct a solution to this challenging problem.

Many researchers studied developing robust building detection algorithms. Mayer [7] provides an excellent overview of building detection studies developed between 1984 and 1998. These studies generally use gray-scale aerial images and detect buildings by thresholding and edge detection techniques. Ünsalan and Boyer [13] provided an extended overview as well as a novel method to detect houses and street segments in residential regions. Levitt and Aghdasi [5] used the region growing algorithm on grayscale images and segmented most of the buildings in their test images. They verified the existence of buildings using the edge map.

Some researchers used the shadow information to detect buildings. Huertas and Nevatia [3] used the relationship between buildings and shadows. They first extracted corners from the image. They labeled these corners as either bright or shadow. They used the bright corners to form rectangles. Shadow corners confirm building hypothesis for these rectangles. McKeown *et al.* [4] detected shadows in aerial images by thresholding. They showed that, building shadows and their boundaries contain important information about building heights and roof shapes. Zimmerman [16] integrated basic models and multiple cues for building detection. He used color, texture, edge information, shadow, and elevation data to detect buildings. His algorithm extracts buildings using blob detection. Tsai [11] compared several invariant color spaces including HSI, HSV, HCV, YIQ, and YC_bC_r models for shadow detection and compensation in aerial images. Matsuoka *et al.* [14] also used shadows to model buildings. They showed that shadow information can be used to estimate damages and changes in buildings. Neuenschwander *et al.* [8] and Rüther *et al.* [10] used snake models to determine exact building shapes from the edge information.

In this study, we present a novel method for automatic building detection from color aerial images. We first extract areas of interest using invariant color features. We also extract shadow information using invariant color features. We use shadow segments to determine the illumination direction and verify building locations. Finally, we present a novel method to determine shapes of the buildings using the edge information.

II. Building Detection

We consider color aerial images in RGB format. Our aim is to detect buildings automatically. We provide a sample test image (with its buildings labeled manually) in Fig. 1. To detect buildings similar to the ones given in this image, we benefit from color invariants. We explore them next.

A. Detecting Rooftop and Shadow Pixels

Color invariants help extracting color properties of objects without being affected by imaging conditions. Imaging conditions can be counted as, the illumination of the environment, the surface properties of the object, the highlights or shadows on the object, and the change of the angle of view. Gevers and Smeulders [2] proposed several color invariants.

Fig. 1. Manually labeled buildings from the Istanbul data set.

We extract the color information in the aerial image using a specific color index based on a previous study [12]. There, we used multispectral information from satellite images as the red and near-infrared bands. Here, we follow the same strategy to define a color invariant, but with the red and green bands of the color aerial image as

$$\psi_r = \frac{4}{\pi} \arctan\left(\frac{R-G}{R+G}\right) \tag{1}$$

where R stands for the red band and G stands for the green band of the color image. This color invariant has the value of unity for red colored objects independent of their intensity values. Similarly, it has the value of minus unity for green colored objects in the image. Therefore, the red rooftops (of buildings) can be easily segmented using ψ_r. Since most buildings have red rooftops in the test region, this invariant is of great use to detect buildings. To segment out the red rooftops automatically, we benefit from Otsu's thresholding method [9].

Shadow information is dominantly found in the blue band of the RGB color space [13]. So, we can propose a similar color invariant (to Eqn. 1) to enhance shadow regions as

$$\psi_b = \frac{4}{\pi} \arctan\left(\frac{B-G}{B+G}\right) \tag{2}$$

where B stands for the blue band of the color image. In the ψ_b image, shadow regions have high intensity values. Therefore, they can easily be segmented out by a simple thresholding operation. To segment out the shadow regions automatically, we again benefit from Otsu's thresholding method [9].

B. Estimating the Illumination Direction

In some previous approaches, the illumination direction is provided to the system manually [15]. We assume that the illumination direction can be estimated if a connected rooftop and shadow region couple can be detected in the image. We consider the illumination direction as the direction of the vector beginning from the center of the rooftop region to the center of the shadow region.

For the rooftop and shadow couple, if center of the rooftop region is at (x_b, y_b), and the center of the shadow region is at (x_s, y_s), then the illumination angle is

$$\theta = \arctan\left(\frac{|y_b - y_s|}{|x_b - x_s|}\right) \tag{3}$$

The quadrant θ lies is also important. We can adjust θ according to its actual quadrant as

$$\theta = \begin{cases} \theta & \text{if} \quad x_s > x_b, y_s < y_b \\ \pi - \theta & \text{if} \quad x_s < x_b, y_s < y_b \\ \pi + \theta & \text{if} \quad x_s < x_b, y_s > y_b \\ 2\pi - \theta & \text{if} \quad x_s > x_b, y_s > y_b \end{cases} \tag{4}$$

C. Verifying the Building Appearance

If the rooftop (of a building) to be detected is not red, then our color invariant ψ_r may not be sufficient to detect it from the aerial image. To detect such rooftops (hence buildings), we have to look for other cues. Since we determined the illumination angle θ in Eqn. 4, this information may be of help to verify red rooftops as well as infer non-red rooftops. To do so, we introduce a hypothesis test such that; if we detect a shadow somewhere in the image it should originate from a building. Therefore, we check for the possible locations of a building based on the illumination direction and the center position of the shadow region as

$$(x_e, y_e) = (x_s + d\cos\theta, y_s + d\sin\theta) \tag{5}$$

where (x_e, y_e) represents the coordinates of the estimated building center. d is the possible distance that a building can be located. In this study, we use this distance as 17 pixels considering the size of buildings in our test set. We provide a simple illustrative example for building verification in Fig. 2.

In Fig. 2, the first image shows a sample building without a red rooftop. In the second image, the dark region represents a shadow segment that is extracted using the color invariant ψ_b. Since the building rooftop is not red, the building location could not be detected by thresholding ψ_r. Therefore, we use the illumination direction information to estimate the possible building location. The illumination angle and direction is calculated using the red rooftop and shadow couples of other buildings (in the image) as we described in the previous section. We provide this result next. We assume that, the building should be in the opposite direction of the illumination vector. We illustrate building center estimation process in the third image. We locate a 30×30 window on this center coordinate. Hereafter, we will call this region as the estimated building segment.

III. Determining Building Shapes

Most buildings have rectangular shapes. To use this property, we introduce a novel box-fitting approach. We benefit from edges of buildings in this method. We use the Canny edge detector for extracting edges of candidate building regions [1]. From edges, the rectangle that represents the building will be constructed. The proposed algorithm works on both the red and non-red rooftop building segments. Therefore, our method does not specifically depend on the red rooftop information.

(a) A non-red rooftop building

(b) The detected shadow segment and illumination direction

(c) Possible location of the building

Fig. 2. Using the shadow information to detect non-red rooftop buildings

We introduce the building shape determination process for two purposes. First, presenting a reliable shape of a building is important for constructing a land map. Second, we verify the appearance of buildings in estimated building segments (which are detected by using only shadow and illumination information).

The proposed box-fitting method discards edges not belonging to building segments. It processes edges of one segment at a time. For each segment, we detect possible corners of line segments it contains. We calculate the angle between lines for each corner. We call this angle as β. We choose the corner which has the smallest $|\pi/2-\beta|$ value and satisfies $|\pi/2-\beta| < \beta_{err}$ condition as a starting point. Due to noise and the viewpoint of the aerial camera, corners may not be exactly $\pi/2$ radians apart. We use a tolerance $\beta_{err} = 0.05\pi$ radians in our experiments to compensate this. We put the initial seed box on the corner which is chosen as the starting point. Then, the box expands in the opposite side of the starting point. We stop the iteration process when the energy E of the box reaches a minimum. Energy of the box is defined as the sum of minimum distances between building edge pixels and the box contour pixels as

$$E = \sum_{i=1}^{n} min(\sqrt{(x_v(i) - x_e(j))^2 - (y_v(i) - y_e(j))^2}) \quad (6)$$

where, $(x_v(i), y_v(i))$ represents coordinates of ith pixel on the contour of the box. $(x_e(j), y_e(j))$ represents the jth pixel of the building edges.

Since the buildings we are looking for are constructed by rectangular shaped structures, it makes sense to try to fit a rectangular shape to the detected edges. The algorithm gives successful results even the edges are not well-determined, there is no closed shape, or the corners are not found. The algorithm also needs less computation time compared to the active contour methods. We provide an example to this box fitting method in Fig. 3.

In Fig. 3, the first image represents edges that belong to a sample building segment. In the second and third images, we demonstrate the box-fitting approach. As can be seen, an initial seed box is located on the corner which has smaller $|\pi/2 - \beta|$ angle between its edges. Then we expand the box outwards until the smallest box energy (E) is achieved. In the third image, the final building shape is provided.

We provide a real example to this box fitting method next. We continue from the non-red rooftop building given in Fig. 2. We provide the final extracted shape of this building in Fig. 4.

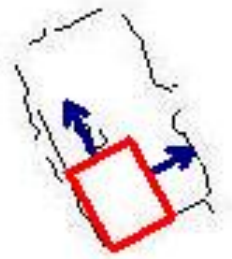
As an outcome of this algorithm, we reject the building segment if the box can not fit the shape in this region. Therefore, we also verify the appearance of buildings in the estimated building segments using this algorithm.

IV. Experimental Results

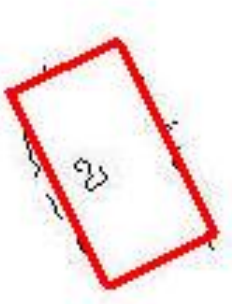
We tested our building detection system on 12 aerial images. They have different illumination conditions. All test images are in RGB color format with a $0.3\ m/pixel$ resolution. Test images are taken from Istanbul, where buildings generally have red rooftops. Some building detection results can be seen in Fig. 5. These examples indicate that our method works fairly well for the test images.

We next tabulate our building detection performance using Lin and Nevatia's formulas [6]. They define the detection performance, DP, and branching factor, BF, for reporting building detection performances. Many researchers use these definitions to report the efficiency of their system. These are

(a) Edges of the building



(b) Initial box position



(c) The fitted box

Fig. 3. Box fitting method to detect the shape of the building.

Fig. 4. An example of the box fitting method to detect the shape of the building.

Fig. 5. Building detection examples.

$$DP = \frac{TP}{TP + TN} \times 100 \quad (7)$$

$$BF = \frac{FP}{TP + FP} \times 100 \quad (8)$$

where TP (true positive) is the number of buildings detected both manually and the automatic approach. FP (false positive) is the number of buildings detected by the automatic approach but not manually. TN (true negative) is the number of buildings detected manually but not by the automatic approach. A building is assumed detected, if at least a small part of it is detected by the automatic approach. The detection percentage, DP, describes how many of the existing buildings in the scene are detected by the automatic approach. The branching factor, BF, gives a hint on how many false alarms are there. The goal is to maximize the DP, while keeping the BF as low as possible.

In Table I, we tabulate the test results for each image in terms of TP and FP. As can be seen, 162 of the 177 buildings are correctly detected in test images. Unfortunately, some tree clusters in the images are detected as possible buildings by our system. They are possible causes of 25 false positives. As a result, we obtain DP as $\mathbf{91,5}\%$ and BF as $\mathbf{13}\%$ for 177 test buildings. Considering this high detection percentage, we can claim that the proposed algorithm is successful.

TABLE I
EVALUATION OF THE DETECTION RESULTS IN THE TEST IMAGE SET

Test image	# buildings	TP	FP
Ist_1	14	14	0
Ist_2	8	8	0
Ist_3	3	3	0
Ist_4	37	34	12
Ist_5	29	28	0
Ist_6	7	7	0
Ist_7	13	9	4
Ist_8	20	17	6
Ist_9	10	10	0
Ist_{10}	24	20	3
Ist_{11}	6	6	0
Ist_{12}	6	6	0
Total	177	162	25

For a 500×500 pixel image which has 15 buildings, the total building detection time is about **10 sec**. This detection time is calculated on a typical desktop computer. The codes are written in Matlab. We can see that, the time needed for our building detection system is also fairly low.

A. Validity of Results

Our method gives reliable results for aerial images having buildings with red and non-red rooftops. If there are non-red rooftop buildings in the test image, then at least one red rooftop and shadow couple should be detected to determine the illumination angle. After determining the illumination angle, locations of buildings that have different rooftop colors can be estimated by our method.

V. CONCLUSIONS

We present a method to detect buildings (via their rooftops) in color aerial images. We use color invariant features to extract red rooftops and shadows. We verify the appearance of buildings (having red and non-red rooftops) using the shadow information. Then we improve on the recognition capability by searching for buildings in the neighborhood of shadow segments. Finally, we determine building shapes by our novel box fitting algorithm. Experimental results indicate the high performance of our building detection method on a diverse data set. We plan to extend this method to L and U shaped buildings in a future study.

REFERENCES

[1] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, November 1986.

[2] T. Gevers and A. W. M. Smeulders, "Pictoseek: Combining color and shape invariant features for image retrieval," *IEEE Transactions on Image Processing*, pp. 102–119, 2000.

[3] A. Huertas and R. Nevatia, "Detecting buildings in aerial images," *Computer Vision, Graphics and Image Processing*, vol. 41, pp. 131–152, 1988.

[4] R. B. Irvin and D. M. McKeown, "Methods for exploiting the relationship between buildings and their shadows in aerial imagery," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 19, no. 1, pp. 1564–1575, 1989.

[5] S. Levitt and F. Aghdasi, "Fuzzy representation and grouping in building detection," in *International Conference on Image Processing*, vol. 3, 2000, pp. 324–327.

[6] C. Lin and R. Nevatia, "Building detection and description from a single intensity image," *Computer Vision and Image Understanding*, vol. 72, no. 2, pp. 101–121, November 1998.

[7] H. Mayer, "Automatic object extraction from aerial imagery - A survey focusing on buildings," *Computer Vision and Image Understanding*, vol. 74, pp. 138–149, 1999.

[8] W. Neuenschwander, P. Fua, G. Székely, and O. Kübler, *Automatic extraction of man-made objects from aerial and space images*. Birkhäuser Verlang, 1995.

[9] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans on SMC*, vol. 9, pp. 62–66, 1979.

[10] H. Ruther, H. M. Martine, and E. G. Mtalo, "Application of snakes and dynamic programming optimization technique in modelling of buildings in informal settlement areas," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 56, no. 4, pp. 269–282, July 2002.

[11] V. J. D. Tsai, "A comparative study on shadow compensation of color aerial images in invariant color models," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 6, pp. 1661–1671, 2006.

[12] C. Ünsalan and K. L. Boyer, "Linearized vegetation indices based on a formal statistical framework," *IEEE Trans on GeoRS*, vol. 42, pp. 1575–1585, 2004.

[13] ——, "A system to detect houses and residential street networks in multispectral satellite images," *Computer Vision and Image Understanding*, vol. 98, pp. 432–461, 2005.

[14] T. Vu, M. Matsouka, and F. Yamazaki, "Shadow analysis in assisting damage detection due to earthquake from quickbird imagery," *Proceedings of the 10th international society for photogrammetry and remote sensing congress*, pp. 607–611, 2004.

[15] G. Zhou, W. Chen, J. Kelmelis, and D. Zhang, "A comprehensive study on urban true orthorectification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 9, pp. 2138–2147, 2005.

[16] P. Zimmermann, "A new framework for automatic building detection analyzing multiple cue data," in *IAPRS*, vol. 33, 2000, pp. 1063–1070.