

ORIENTATION BASED BUILDING OUTLINE EXTRACTION IN AERIAL IMAGES

Andrea Kovacs^{a,b} and Tamas Sziranyi^b

^a Pazmany Peter Catholic University, Prater 50/A, 1083, Budapest, Hungary

^b MTA SZTAKI, Kende 13-17, 1111, Budapest, Hungary
andrea.kovacs@sztaki.hu, sziranyi@sztaki.hu

Commission VII, WG VII/4

KEY WORDS: Orientation Estimation, Building Detection, Feature Extraction, Aerial Image Detection, Urban

ABSTRACT:

The goal of this paper is to extract automatically the building contours regardless of shape. By extracting these contours, detection results will be more accurate, giving useful information about urban area, which is important for many tasks, like map updating and disaster management. First, we extract local feature points from the image, based on a modification of Harris detector's saliency function, which can represent urban area and building effectively. This point set is then used to define the main orientation of the buildings, which characterizes well an urban region and helps to define directions, where object contours have to be searched. Second, we applied shearlet approach to extract edges in the defined directions. This results in an edge map, which helps us to determine point subsets belonging to the same building. Convex hulls of the point subsets is used for contour initialization, then region based Chan-Vese Active Contour method is applied to extract the accurate building outlines.

1 INTRODUCTION

Automatic evaluation of aerial photographs is a very important research topic, as the manual analysis is very time consuming. Nowadays, there are many approaches for multispectral or synthetic aperture radar (SAR) images, but developing methods handling optical aerial or satellite images have also big importance (Zhong and Wang, 2007), (Dey et al., 2011). When working on optical photographs, the challenge is the large variety of features: images can be grayscale or containing poor color information, scanned in different seasons and in altering lighting conditions. In this case, pixel neighborhood processing techniques like multi-layer difference image or background modeling (Benedek and Szirányi, 2008) cannot be adopted efficiently since details are not comparable.

In our work, we concentrate on building detection, which is a very important task, as land area might be changing dynamically and a continuous periodic administration is necessary to have up-to-date information. This is very useful for urban development analysis, map updating and also helps in disaster management, vegetation monitoring and discovering illegal surface forming activities. The challenges of building detection partly come from the aforementioned diverse imaging circumstances, causing different color, contrast and shadow conditions. On the other hand, the shape of different buildings is quite various, which needs sophisticated techniques to have more accurate results.

There is a wide range of publications in remote sensing topic for urban area and building detection. However, we concentrate on novel approaches which will be used for comparison later. State-of-the-art building detection approaches can be divided into two main groups. The first group contains methods which only localize buildings without any shape information (Sirmacek and Unsalan, 2009) and (Sirmaçek and Ünsalan, 2011). In these approaches, only the location of the building is detected. In (Sirmacek and Unsalan, 2009) a SIFT salient point based approach is introduced for urban area and building detection (denoted by SIFT-graph in the experimental part). This method uses two templates (a light and dark one) for detecting buildings. After extracting feature points representing buildings, graph based tech-

niques are used to detect urban area. The given templates help to divide the point set into separate building subsets, then the location is defined. However, in many cases, the buildings cannot be represented by such templates, moreover sometimes it is hard to distinguish them from the background based on the given features. (Sirmaçek and Ünsalan, 2011) proposes a method to detect building positions in aerial and satellite images based on Gabor filters (marked as Gabor filters in the experimental part), where different local feature vectors are used to localize buildings with data and decision fusion techniques.

The other group contains approaches which use some shape templates (e.g. rectangles) for detecting the buildings (Song et al., 2006), (Sirmaçek and Ünsalan, 2008) and (Benedek et al., 2012). In this case, beside the location, additional information is given about the size, orientation and shape. In (Song et al., 2006) a segment-merge technique is introduced (Segment-Merge), which represents a distinct trend. This method considers building detection task as a region level problem and assumes that buildings are homogeneous areas (either regarding color or texture information), and based on this fact, they can be distinguished from the background. In the first step, the background is subtracted, then some shape and size constraints are created to define building objects. However, the basic assumptions influences the success of the approach: sometimes buildings cannot be distinguished from the background effectively by using color and texture features, therefore the further steps will also fail.

(Sirmaçek and Ünsalan, 2008) (named as Features-Canny in the experimental part) combines roof color, shadow and edge information in a two-step process. First, a built-in candidate is defined based on color and shadow feature, then a rectangle template is fitted using a Canny edge map. This sequential method is very sensitive to the deficiencies of both steps: the inappropriate shadow and color information results in false candidates, and accurate detection is not possible with a malfunctioning edge map.

A novel building detection approach is introduced in (Benedek et al., 2012), using a global optimization process, considering observed data, prior knowledge and interactions between the neigh-

boring building parts (marked later as bMBD). The method uses low-level (like gradient orientation, roof color, shadow, roof homogeneity) features which are then integrated to have object-level features. After having object (building part) candidates, a configuration energy is defined based on a data term (integrating the object-level features) and a prior term, handling the interactions of neighboring objects and penalizing the overlap between them. The optimization process is then performed by a bi-layer multiple birth and death optimization.

Although the second group provides some information about the shape, it is still just an approximation. Therefore, our aim is to construct a method, which can deal with the shape diversity. In the first step, we generate a feature point set, based on our modification of the Harris corner detector (Kovács and Szirányi, 2012), which is able to represent object contours effectively. Next, we calculate main directions in the surroundings of feature points to get orientation information, which characterizes the urban area as well. Then, an improved edge map is constructed by strengthening the edges only in the calculated main directions with shearlet method (Easley et al., 2009). Based on the edge map, the feature point set is divided into subsets by a graph based connectivity detection (a similar was introduced in (Sirmacek and Unsalan, 2009)), where each subset represents a building candidate. Using the convex hull of the point subset for initialization, Chan-Vese active contour method (Chan and Vese, 2001) detects the final boundary of the building. Evaluating the method on aerial images provided by the Hungarian Institute of Geodesy, Cartography and Remote Sensing, the initial results show that our proposed method is able to detect buildings more efficiently and it can be a rival for other state-of-the-art methods.

2 FEATURE POINT EXTRACTION

2.1 Original Harris detector

The detector was introduced in 1988 (Harris and Stephens, 1988) and based on the principle that at corner points intensity values change largely in multiple directions. By considering a local window in the image and determining the average changes of image intensity result from shifting the window by a small amount in various directions, all the shifts will result in large change in case of a corner point. Thus corner can be detected by finding when the minimum change produced by any of shifts is large.

The method first computes the Harris matrix (M) for each pixel in the image, consisting the product of the first order derivatives smoothed by a Gaussian window. Then, instead of computing the eigenvalues of M , an R corner response is defined:

$$R = \text{Det}(M) - k * \text{Tr}^2(M), \quad (1)$$

with Det and Tr denoting the determinant and trace of M and k is a coefficient, usually around 0.04.

This R characteristic function is used to detect corners. R is large and positive in corner regions, and negative in edge regions (Figure 1(b)). By searching for local maxima of R , the Harris keypoints can be found (see Figure 1(c)).

2.2 Modified Harris based feature map

When working on contour detection, contour points have to be emphasized with some techniques. In our previous work (Kovács and Szirányi, 2012), we introduced a modification of the original Harris method, which is able to emphasize edges and corners

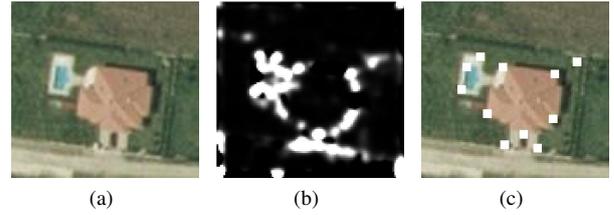


Figure 1: Operation of the Harris detector: (a) shows the original image; (b) is the R characteristic function; (c): Keypoints chosen as local maxima of R .

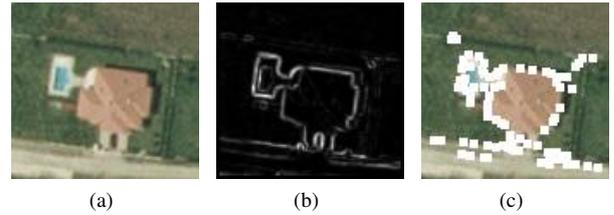


Figure 2: Operation of the modified Harris detector: (a) shows the original image; (b) is the proposed $R_{\log\max}$ function; (c): Keypoints chosen as local maxima of $R_{\log\max}$.

equally, therefore can be applied efficiently for generating a feature map for active contour approaches. The proposed modification looks as follows:

$$R_{\log\max} = \max(0, \log[\max(\lambda_1, \lambda_2)]), \quad (2)$$

where λ_1 and λ_2 denote the eigenvalues of M . When emphasizing corners and edges, they both have one large component, thus $\max(\lambda_1, \lambda_2)$ function separates the flat and non-flat regions accurately. To produce a steady feature map, the dynamics of the characteristic function should be compressed into a balanced distribution by keeping the necessary strength of the main attractors. The natural logarithmic (\log) function satisfies this condition: it has a balanced output for both corner and edge saliency. The target set of the $R_{\log\max}$ is the positive domain (when it is used as a feature map), thus the outer max function is responsible for replacing negative values of small λ (points in flat regions) with zeros.

By calculating the local maxima of the proposed $R_{\log\max}$ function, the modified feature point set is defined, see Figure 2.

2.3 Urban area detection

The advantage of the (original and modified) Harris detector is its strong invariance to rotation, illumination variation, image noise and robustness on fixed scales. Therefore it can be an efficient tool in aerial image segmentation and can handle the altering characteristics of different images. The first question in applying the proposed function for the building detection task was if the proposed extended feature point set could represent the urban area. Is it possible to use these points for building detection? To get answers to these questions, we have evaluated the point set for urban area detection. For testing, we used spatial voting, which was proposed in (Sirmacek and Unsalan, 2010) for Gabor filter based local feature points. The method assumes that when detecting an urban area, many local feature points should be in it located closely in the spatial domain, and around the points there is a high possibility of urban features. Therefore, the constructed voting matrix has the highest vote at the location of the feature

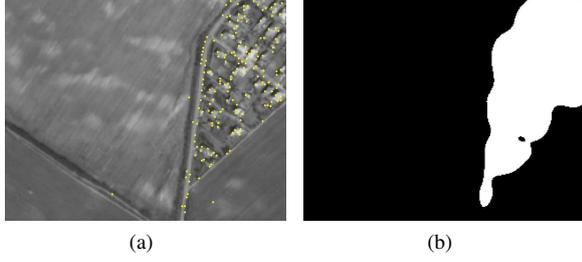


Figure 3: Urban area detection with spatial voting: (a) shows the original image; (b) is the detected urban area.

point (x_i, y_i) for the i^{th} point), and the vote is decreasing around it in accordance with the spatial distance:

$$V(x, y) = \sum_{i=1}^K \frac{1}{2\pi\sigma_i^2} \exp\left(-\frac{(x-x_i)^2 + (y-y_i)^2}{2\sigma_i^2}\right), \quad (3)$$

where σ_i is the parameter for voting proximity for point (x_i, y_i) . After calculating V for every pixel in the image, Otsu thresholding (Otsu, 1979) was applied to distinguish urban area from background.

Testing the method for modified Harris feature points for many different aerial images, the results showed that urban area is well represented by this point set. A typical result can be seen in Figure 3, where a grayscale image was used with limited color information and some cloud shadows generating false contours in the left size of the image, which makes the detection harder. However, the generated point set is mostly situated in the urban area and the spatial voting can detect the area quite accurately.

3 ORIENTATION BASED BUILDING DETECTION

Our tests have shown that the proposed Harris point set is able to represent urban area, therefore features of the points can also be used to extract information of the urban area to improve building detection. Our idea was to extract the main direction in the small neighborhood of the feature points which can characterizes the urban area. Edges in the defined direction are strengthened with the shearlet method (Yi et al., 2009). Building candidates are determined then based on the improved edge map and the feature point subset. The accurate building contours are generated with Chan-Vese active contour method (Chan and Vese, 2001).

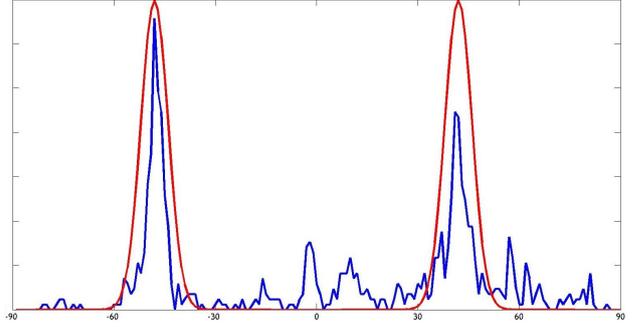
3.1 Orientation estimation

A small urban area has buildings with connected orientation. In most cases, houses are oriented according to some bigger structure (e. g. the road network), therefore main orientation of the area can be defined. As the proposed modified Harris point set represents the area, our idea was to calculate the main direction of the buildings of the area based on this point set. (Benedek et al., 2012) used a low level feature, called local gradient orientation density, where the surroundings of a pixel was investigated whether it has perpendicular edges or not. Now we use a similar feature to calculate the main direction of a feature point's neighborhood. Let us denote the gradient vector by ∇g_i with $\|\nabla g_i\|$ magnitude and φ_i^∇ orientation for the i^{th} point. By defining the $n \times n$ neighborhood of the point with $W_n(i)$ (where n depends on the resolution), the weighted density of φ_i^∇ is as follows:

$$\lambda_i(\varphi) = \frac{1}{N_i} \sum_{r \in W_n(i)} \frac{1}{h} \cdot \|\nabla g_r\| \cdot k\left(\frac{\varphi - \varphi_r^\nabla}{h}\right), \quad (4)$$



(a)



(b)

Figure 4: Orientation estimation for an image: (a) shows the feature points in yellow; (b) shows the $\lambda_i(\varphi)$ orientation density function of the points in blue, calculated for 15×15 neighborhood. $\varphi \in [-90, +90]$ is the horizontal axis, the number of points is the vertical axis. The $\eta_2(\cdot)$ two-component Gaussian mixture is in red, detected peaks are $\theta = -47$ and $\theta_{\text{ortho}} = +53$.

with $N_i = \sum_{r \in W_n(i)} \|\nabla g_r\|$ and $k(\cdot)$ kernel function with h bandwidth parameter.

We define a main orientation for (i^{th}) feature point as:

$$\varphi_i = \underset{\varphi \in [-90, +90]}{\operatorname{argmax}} \{\lambda_i\}. \quad (5)$$

After calculating the direction for all the K feature points, we define the density function ϑ of their orientation:

$$\vartheta(\varphi) = \frac{1}{K} \sum_{i=1}^K H_i(\varphi), \quad (6)$$

where $H_i(\varphi)$ is a logical function:

$$H_i(\varphi) = \begin{cases} 1, & \text{if } \varphi_i = \varphi \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

We expect that the density function ϑ will have two main peaks (because of the perpendicular edges of buildings), see Figure 4. This can be measured by correlating ϑ to a bimodal density function:

$$\alpha(m) = \int \vartheta(\varphi) \eta_2(\varphi, m, d_\vartheta) d\varphi, \quad (8)$$

where $\eta_2(\cdot)$ is a two-component mixture of Gaussian, with m and $m + 90$ mean values and d_ϑ is standard deviation for both components. The value θ of the maximal correlation can be obtained as:

$$\theta = \underset{m \in [-90, +90]}{\operatorname{argmax}} \{\alpha(m)\}. \quad (9)$$

And the corresponding orthogonal direction (the other peak of $\alpha(m)$):

$$\theta_{\text{ortho}} = \begin{cases} \theta - 90, & \text{if } \theta \geq 0 \\ \theta + 90, & \text{otherwise} \end{cases} \quad (10)$$

Thus, we expect building edges to be in the calculated main orientation and we try to enhance edges in the given directions. We tested the orientation estimation approach for different datasets and different window size ($n \times n$ neighborhood). The results show that with larger size $n = 15$, the density function is smoother (see Figure 4(b)), while with smaller size ($n = 5$) rougher and blurred, but the main characteristics and the main peaks were obvious in both cases, therefore we only showed the result of 15×15 neighborhood in Figure 4.

3.2 Edge detection with shearlet transform

Now, the main orientations have been defined. The next step is to enhance edges in these directions to extract useful edge information from the image that can be later combined efficiently with the feature point set. There are different approaches which uses directional information like Canny edge detection (Canny, 1986) using the gradient orientation; or (Perona, 1998) which is based on anisotropic diffusion, but cannot handle the situation of multiple orientations (like corners). Other single orientation methods exist, like (Mester, 2000) and (Bigun et al., 1991), but the main problem with these methods is that they calculate orientation in pixel-level and lose the scaling nature of orientation, therefore they cannot be used for edge detection. In our case we need to enhance edges constructed by joint pixels, thus we searched for such edge detection method which can handle orientation as well. Moreover, as searching for building contours, the algorithm must handle corner points as well. Shearlet transform (Yi et al., 2009) has been lately introduced for efficient edge detection, as unlike wavelets, shearlets are theoretically optimal in representing images with edges and, in particular, have the ability to fully capture directional and other geometrical features.

For an image u , the shearlet transform is a mapping:

$$u \rightarrow SH_{\psi}u(a, s, x), \quad (11)$$

providing a directional scale-space decomposition of u with $a > 0$ is the scale, s is the orientation and x is the location:

$$SH_{\psi}u(a, s, x) = \int u(y)\psi_{as}(x - y)dy = u * \psi_{as}(x), \quad (12)$$

where ψ_{as} are well localized waveforms at various scales and orientations. As we are working with a discrete transform, a discrete set of possible orientations is used, for example $s = 1, \dots, 16$. In our case, the main orientation of the image θ is calculated (see Section 3.1), therefore our aim is to strengthen the components in the given direction on different scales as we only want to detect edges in the main orientation. The first step is define the s subband which includes θ and θ_{ortho} :

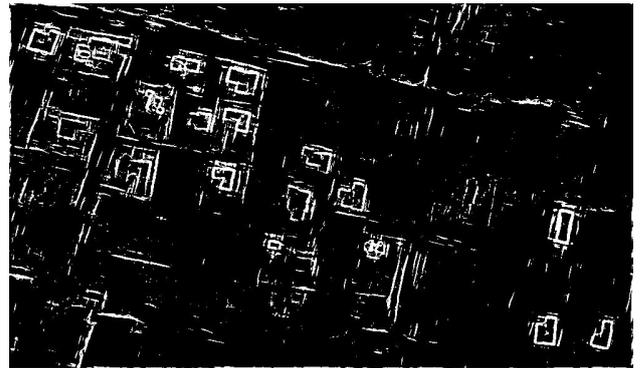
$$\begin{aligned} \tilde{s} &= \left\{ s_i : (i-1)\frac{2\pi}{s} < \theta \leq i\frac{2\pi}{s} \right\}, \\ \tilde{s}_{ortho} &= \left\{ s_j : (j-1)\frac{2\pi}{s} < \theta_{ortho} \leq j\frac{2\pi}{s} \right\}. \end{aligned} \quad (13)$$

After this, the $SH_{\psi}u(a, \tilde{s}, x)$ and $SH_{\psi}u(a, \tilde{s}_{ortho}, x)$ subbands have to be strengthened. For this reason, the weak edges (values) have been eliminated with a hard threshold and only the strong coefficients are amplified.

Finally, the shearlet transform is applied backward (see Eq.12) to get the reconstructed image, which will have strengthened edges in the main directions. The strengthened edges can be easily detected with Otsu thresholding, results can be seen in Figure 5. While the pure Canny method detects the edges sometimes with discontinuities, the shearlet based edge strengthening helps to eliminate these problems.



(a)



(b)

Figure 5: Comparing the edge maps for u^* channel : (a) shows the result of the pure Canny edge detection; (b) is the result of the shearlet based edge strengthening.

As building colours may vary largely and the shadow effect have to be reduced to eliminate false contours when detecting buildings, we used two color channels in the edge strengthening step: for red buildings we used the u^* component of CIE L^*u^*v advised in (Muller and Zaum, 2005); for grey buildings we applied the C_b component of YC_bC_r space, which was found to separate grey coloured objects and their shadows the most effectively in (Tsai, 2006). First, we extracted the red building outlines based on the u^* edge map, after that the remaining buildings were detected in the C_b -based map.

3.3 Building contour detection

After defining the main orientation for the extended Harris point set for the urban region, the shearlet based edge strengthening approach enhanced the edges in the main directions, resulting in an edge map S . In the next step we will compound the feature point set and edge map with a graph based representation, which was introduced in our previous work (Kovacs and Szirányi, 2010); based on the generated edge map, connected feature point subgraphs are determined, indicating building candidates. The E edge network of $G = (V, E)$ graph is constructed by connecting $v_i = (x_i, y_i)$ and $v_j = (x_j, y_j)$, the i^{th} and j^{th} vertices of the V feature point set, if they satisfy the following conditions:

1. $S_{(x_i, y_i)} = 1$,
2. $S_{(x_j, y_j)} = 1$,
3. \exists a finite path between v_i and v_j in S .

The result of this a graph composed of many separate subgraphs, where each subgraph indicates a building candidate. However,

there might be some singular points and some smaller subgraphs (points and edges connecting them) indicating noise. To discard them, we select subgraphs having points over a given threshold.

To detect the accurate contour of the buildings, we use Chan-Vese active contour algorithm (Chan and Vese, 2001) and initialize the contour for a building candidate as the convex hull of the vertices of the subgraph.

Main directional edge emphasis may also enhance road and vegetation contours, moreover some feature points can also be located on these edges. Therefore after the contour extraction step the results have to be supervised to filter out misdetections. When detecting false objects, like road parts or land section borders, the edges in the detected area are unidirectional, unlike buildings, which have either orthogonal or multidirectional contours. Thus, the directional distribution of edges is evaluated in the extracted area (see the technique in Section 3.1) and unidirectional hits are eliminated. Here, we use again the correlation to a bimodal density function (Eq. 8) then measure and threshold the α value to select multidirectional hits.

Figure 6 shows the result of the building detection with the detected and filtered contours. Based on the contours, we can estimate the location of the buildings which will be useful in the further work for evaluation and comparison.

4 EXPERIMENTS

We have evaluated our proposed method for the Szada dataset provided by the Hungarian Institute of Geodesy, Cartography and Remote Sensing. This dataset was also used in (Benedek et al., 2012) for evaluation and for comparison with different methods ((Sirmacek and Unsalan, 2009), (Sirmacek and Ünsalan, 2011), (Song et al., 2006) and (Sirmacek and Ünsalan, 2008)), therefore quantitative test results are available. In Table 1 the quantitative results for Szada dataset is shown. The name of the compared methods is abbreviated as marked in the introduction part. The complete dataset contains 57 buildings out of which our method is able to detect 55 buildings (meaning 2 misdetections) with 0 false positive object. In this case we used the location of the buildings (see Figure 6(b)) for evaluation. By comparing this with the other approaches, one can see that our method is able to outperform the others.

For qualitative evaluation we used the detected outlines (see Figure 6(a)). A part of this image was cut and enlarged to compare our method qualitatively with (Benedek et al., 2012). Figure 7 shows the detailed differences. However, rectangular templates provide a very close estimation for the shape of the buildings, the fine details are lost. Unlike shape templates, active contour based techniques do not apply any restrictions for the shape and able to detect the varying contour parts more accurately.

Although active contours are able to cope with the altering shapes, sometimes they suffer from the lack of contrast difference between the building and the background and have difficulties when detecting contours (like missing a part of the building outline, see the building in the bottom-middle of Figure 6(a)). Moreover, sometimes buildings are oriented variously and the edge strengthening step misses their edges, which may result in misdetections.

5 CONCLUSION AND FUTURE WORK

In this paper, we have introduced a novel orientation based method for building detection in aerial images. The proposed method



(a)



(b)

Figure 6: Result of the building detection: (a) shows the detected contours; (b) is the estimated locations of the detected buildings.



(a)

(b)

(c)

Figure 7: Qualitative comparison of MPP-based and proposed method: (a) is the original image part; (b) shows the result of MPP-based method; (c) is the result of the proposed approach.

first calculates a feature point set based on the modification of the well-known Harris corner detector. In the first step, we have proven the detector's ability to characterize urban area by testing the point set with a voting matrix technique for urban area detection. The orientation feature of the point set is then used to define main direction of the urban area, to make an edge strengthening in the given directions. Shearlet transform have been applied for the orientation based edge enhancing, as it is able to handle orientation information, even in multidirectional cases (like corners). The improved edge information is combined with the feature point set and a graph based technique was introduced to get feature point subgraphs as building candidates. Finally, Chan-Vese nonparametric active contour approach was applied to extract the building contours. The proposed method have been compared with other state-of-the-art algorithms quantitatively and qualitatively. The results showed that our approach can detect buildings more effectively than the others, either for pure localization and contour extraction.

However, this paper is just the first step of work, introducing the main principles and needs more evaluation for different databases.

Szada dataset (57 buildings)	Missing objects	False objects
SIFT-graph (Sirmacek and Unsalan, 2009)	17	26
Gabor filter(Sirmaçek and Ünsalan, 2011)	17	23
Features-Canny (Sirmaçek and Ünsalan, 2008)	10	18
Segment-Merge (Song et al., 2006)	11	5
bMBD (Benedek et al., 2012)	4	1
Proposed	2	0

Table 1: Quantitative results for Szada dataset

Moreover, there are still some open questions, which have to be answered in the near future. If the resolution of the aerial image is smaller, with larger scanned urban area, the orientation of the buildings may vary, which have to be handled, for example by correlating the orientation distribution function with multiple bimodal Gaussian functions. Furthermore, the active contour method might suffer from difficulties when building contours are hidden by other structures like trees. In this case some prior constraints (like edge parts running in the defined main orientations) can be introduced.

ACKNOWLEDGMENT

This work was supported by the Hungarian Scientific Research Fund (OTKA) under grant number 76159.

REFERENCES

- Benedek, C. and Szirányi, T., 2008. Bayesian foreground and shadow detection in uncertain frame rate surveillance videos. *IEEE Trans. Image Processing* 17(4), pp. 608–621.
- Benedek, C., Descombes, X. and Zerubia, J., 2012. Building development monitoring in multitemporal remotely sensed image pairs with stochastic birth-death dynamics. *IEEE Trans. Pattern Analysis and Machine Intelligence* 34(1), pp. 33–50.
- Bigun, J., Granlund, G. H. and Wiklund, J., 1991. Multidimensional orientation estimation with applications to texture analysis and optical flow. *IEEE Trans. Pattern Analysis and Machine Intelligence* 13(8), pp. 775–790.
- Canny, J., 1986. A computational approach to edge detection. *IEEE Trans. Pattern Analysis and Machine Intelligence* 8(6), pp. 679–698.
- Chan, T. F. and Vese, L. A., 2001. Active contours without edges. *IEEE Trans. Image Processing* 10(2), pp. 266–277.
- Dey, V., Zhang, Y. and Zhong, M., 2011. Building detection from pan-sharpened geocye-1 satellite imagery using context based multi-level image segmentation. In: *International Symposium on Image and Data Fusion (ISIDF)*, pp. 1–4.
- Easley, G. R., Labate, D. and Colonna, F., 2009. Shearlet-based total variation diffusion for denoising. *IEEE Trans. Image Processing* 18(2), pp. 260–268.
- Harris, C. and Stephens, M., 1988. A combined corner and edge detector. In: *Proceedings of the 4th Alvey VisionConference*, pp. 147–151.
- Kovacs, A. and Szirányi, T., 2010. New saliency point detection and evaluation methods for finding structural differences in remote sensing images of long time-span samples. In: *Advanced Concepts for Intelligent Vision Systems, Springer - Lecture Notes on Computer Sciences, Sydney, Australia*.
- Kovács, A. and Szirányi, T., 2012. Harris function based active contour external force for image segmentation. *Pattern Recognition Letters* 33(9), pp. 1180–1187.
- Mester, R., 2000. Orientation estimation: Conventional techniques and a new non-differential approach. In: *Proc. 10th European Signal Processing Conference*.
- Muller, S. and Zaum, D., 2005. Robust building detection in aerial images. In: *CMRT, Vienna, Austria*, pp. 143–148.
- Otsu, N., 1979. A threshold selection method from gray-level histograms. *IEEE Trans. Systems, Man and Cybernetics* 9(1), pp. 62–66.
- Perona, P., 1998. Orientation diffusion. *IEEE Trans. Image Processing* 7(3), pp. 457–467.
- Sirmaçek, B. and Ünsalan, C., 2008. Building detection from aerial imagery using invariant color features and shadow information. In: *Int. Symp. on Computer and Information Sciences (ISCIS)*.
- Sirmaçek, B. and Ünsalan, C., 2010. Using local features to measure land development in urban regions. *Pattern Recognition Letters* 31, pp. 1155–1159.
- Sirmaçek, B. and Ünsalan, C., 2011. A probabilistic framework to detect buildings in aerial and satellite images. *IEEE Trans. Geoscience and Remote Sensing* 49(1), pp. 211–221.
- Sirmacek, B. and Unsalan, C., 2009. Urban-area and building detection using sift keypoints and graph theory. *IEEE Trans. Geoscience and Remote Sensing* 47(4), pp. 1156–1167.
- Song, Z., Pan, C. and Yang, Q., 2006. A region-based approach to building detection in densely build-up high resolution satellite image. In: *IEEE International Conference on Image Processing*, pp. 3225–3228.
- Tsai, V., 2006. A comparative study on shadow compensation of color aerial images in invariant color models. *IEEE Trans. Geoscience and Remote Sensing* 44(6), pp. 1661–1671.
- Yi, S., Labate, D., Easley, G. R. and Krim, H., 2009. A shearlet approach to edge analysis and detection. *IEEE Trans. Image Processing* 18(5), pp. 929–941.
- Zhong, P. and Wang, R., 2007. A multiple conditional random fields ensemble model for urban area detection in remote sensing optical images. *IEEE Trans. Geoscience and Remote Sensing* 45(12), pp. 3978–3988.