

SHAPE DETECTION OF STRUCTURAL CHANGES IN LONG TIME-SPAN AERIAL IMAGE SAMPLES BY NEW SALIENCY METHODS

Andrea Kovacs and Tamas Sziranyi

Pazmany Peter Catholic University
Prater 50/A, 1083, Budapest, Hungary
kovan1@digitus.itk.ppke.hu
and

MTA SZTAKI
Kende 13-17, 1111, Budapest, Hungary
sziranyi@sztaki.hu

Commission III/3

KEY WORDS: Change detection, Saliency points, Active Contours, Harris corner function, Aerial photos

ABSTRACT:

The paper presents a novel, theoretically well-based methodology to find changes in remote sensing image series. The proposed method finds changes in images scanned with a long time-interval difference in very different lighting and surface conditions. The presented method is basically an exploitation of the Harris saliency function and its derivatives for finding featuring points among image samples; then a new local descriptor will be demonstrated by generating local active contours; a graph based shape descriptor will be shown based on the saliency points of the difference and in-layer features; finally, we prove the method's capabilities for finding structural changes on remote sensing images.

1 INTRODUCTION

Automatic evaluation of aerial photograph repositories is an important field of research since manual administration is time consuming and cumbersome. Long time-span surveillance or reconnaissance about the same area is a requirement for quick and up-to-date content retrieval. The extraction of long-time changes may facilitate several applications in urban development analysis, disaster protection, agricultural monitoring and detection of illegal surface forming activities. The obtained change map should provide useful information about size, shape, or quantity of the changed areas, which could be applied directly by higher level object analyzer modules (Peng et al., 2008), (Lafarge et al., 2008). While numerous state-of-the art approaches in remote sensing deal with multispectral or synthetic aperture radar (SAR) imagery, the significance of handling optical photographs is also increasing (Zhong and Wang, 2007). When comparing archived optical image collections, the photographs are often grayscale or contain only poor color information. This paper focuses on finding the contours of newly appearing/fading out objects in optical aerial images, which were taken with several years time differences partially in different seasons and in different 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.

These optical image sensors provide limited information and we can only assume to have image repositories which contain geometrically corrected and registered (Shah et al., 2008) grayscale orthophotographs.

One possible approach is the postclassification comparison, which segments the input images into different land-cover classes, obtaining the changes indirectly as regions with different classes in the two image layers (Zhong and Wang, 2007). We follow another methodology, like direct methods (Ghosh et al., 2007), (Wiemker, 1997), where a similarity-feature map is derived from

the input photographs (e.g., a DI), then the feature map is separated into changed and unchanged areas. In our present work the question whether somewhere change is proposed is connected to the question that the given place is a salient one or not. Consequently we should find areas, which have high effect both in the difference image and in the newer image. These areas can define the keypoint candidates, indicating newly born objects.

In our work we mainly focus on new objects (buildings, pools, etc.). There are many difficulties when detecting such objects in airborne images: the illumination and weather circumstances may vary, resulting different colour, contrast and shadow conditions; position may yield different point of views and hidden structures. These objects are quite various, which also makes the detection tough. Our direct method does not use any land-cover class model, and attempts to detect changes which can be discriminated by low-level features. However, our approach is not a pixel-neighborhood MAP system as in (Benedek and Szirányi, 2009), but a connection system of nearby saliency points. These saliency points define a link network by using local graphs for outlining the local hull of the objects. Considering this curve as a starting spline, we search for objects' boundaries by active contour iterations. For the above features the main saliency detector is calculated as a discriminative function among the functions of the different layers. We show that the Harris detector, introduced in 1988 (Harris and Stephens, 1988), is the appropriate function for finding the dissimilarities among different layers, when comparison is not possible because of the different lighting, color and contrast conditions.

Local structure around keypoints is investigated by generating scale and position invariant descriptors, like SIFT. These descriptors describe the local microstructure. However, in several cases a more succinct set of parameters is needed. For this reason we have developed a local active contour ((Xu and Prince, 1997)) based descriptor around keypoints (Kovacs and Sziranyi, 2009). To fit together the definition of keypoints and their active contour around them, we have introduced the Harris corner function as

an outline detector instead of the simple edge functions, gaining a much better characterization of local structure. The generated local contours assigns an individual shape to every keypoint, featured by Fourier descriptors (Licsar and Sziranyi, 2005).

After selecting the saliency points indicating change, we now have to enhance the number of keypoints. Therefore we are looking for saliency points that are not presented in the older image, but exists on the newer one. We call for the Harris corner function again.

Saliency points support the boundary hull definition of objects, constructed by graph based connectivity detection (Sirmacek and Unsalan, 2009) and neighborhood description. This graph based shape descriptor works on the saliency points of the difference and in-layer features. We prove the method in finding structural changes on remote sensing images. We obtain a graph composed of many separate subgraphs. Each of these connected subgraph is supposed to represent an object. However, there might be some unmatched keypoints, indicating noise. The convex hull of the vertices in the subgraphs is applied as the initial contour for running active contour process for the object's boundary.

In the following sections, first we introduce our new change detection procedure (more details about the algorithm can be found in (Kovacs and Sziranyi, 2010b)) by using the Harris function and its derivatives for finding saliency points among image samples. Then a new local descriptor will be demonstrated by generating local active contours (based on (Kovacs and Sziranyi, 2009)). A graph based shape descriptor will be shown based on the saliency points of the difference and in-layer features. Finally, we prove the methods capabilities for finding structural changes on remote sensing images.

2 CHANGE DETECTION WITH HARRIS KEYPOINTS

2.1 Harris corner detector

The detector was introduced by Chris Harris and Mike Stephens in 1988 (Harris and Stephens, 1988). The algorithm 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. Then, instead of computing the eigenvalues of M , an R corner response is defined:

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

This R characteristic function is used to detect corners. R is large and positive in corner regions, and negative in edge regions. By searching for local maximas of R , the Harris keypoints can be found. Figure 1 shows the result of Harris keypoint detection. On Figure 1(b) light regions shows the larger R values, so keypoints will be detected in these areas (Figure 1(c)). Other specific corner/edge detectors can be combined from the eigenvalues of M , as they are compared in (Kovacs and Sziranyi, 2010a).

2.2 Change detection

The advantage of the Harris detector is its strong invariance to rotation, illumination variation, image noise and robustness on

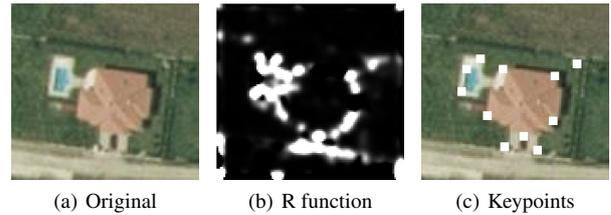


Figure 1: Operation of the Harris detector: Corner points of the original image are chosen as the local maximas of the R characteristic function.



(a) Older image (I_{old})



(b) Newer image (I_{new})

Figure 2: Original image pairs

different scales. Therefore it could be a good feature for detecting changes on airborne images. In this kind of images, changes can mean the appearance of new man-made objects (like buildings or streets), or natural, environmental variations. As image pairs may be taken with large intervals of time, the area may change largely. In our case the parts of the image pairs were taken in 2000 and 2005 (Figure 2). It must be mentioned that these image pairs are registered and represents exactly the same area.

In our work we mainly focus on newly built objects (buildings, pools, etc.). Several difficulties may arise when detecting such objects in airborne images, like the variance of illumination and weather circumstances, resulting in different colour, contrast and shadow conditions. Abundant depth-content (like in urban areas) are disturbed by imaging them from different point of views. Objects can be hidden by other structures like trees, shadows,



Figure 3: Grayscale images generated three different ways: (a) RGB components, (b) R component, (c) U^* component

buildings. These objects are quite various, which also makes the detection tough.

To account the aforementioned difficulties, the first idea is to use some difference of the image pairs. As we are searching for newly-built objects, we need to find buildups that only exist on the newer image (I_{new}), therefore having large effect on the difference image. Our assumption was to find areas which have significant effects both on the difference image and on the newborn image. These areas can define the keypoint candidates, indicating newly made objects.

We need grayscale images both for difference map generating, both for keypoint detection. To be more specific on the objects, the enhancement of change caused by buildups, it is more efficient to apply only the red component of the image (when talking about color images), rather than all components (see Figure 3).

It is worthy to note, that further on the u^* component of the $L^*u^*v^*$ colorspace (Figure 3(c)) will also play an important role in our algorithm.

We examined the intensity based, edge based and R_n based difference map in (Kovacs and Sziranyi, 2010b). Here, R_n marks the normalized characteristic function of Harris detector. Normalization means a simple rescaling of the R function to get positive codomain. Our experiments showed that the R_n based difference map outperforms the two other, therefore it can be used efficiently for keypoint detection.

In the R_n based process we search for such keypoint candidates that accomplish the next two criterias simultaneously:

1. $R_n(I_{diff}) > \epsilon_1$
2. $R_n(I_{new}) > \epsilon_2$

$R_n(I_{diff})$ denotes the difference maps, which is as follows:



Figure 4: Logarithmized difference map

$$I_{diff} = |R_n(I_{new}) - R_n(I_{old})| \quad (2)$$

The first criteria indicates that the candidate has high effect on the difference image, while the second guarantees that it has also big influence on the newer image. ϵ_1 and ϵ_2 are thresholds. It is advised to take smaller ϵ_2 , than ϵ_1 . With this choice the difference map is preferred and has larger weight. Only important corners in the difference map will be marked.

The logarithm of difference map (see comparison in (Kovacs and



(a) Intensity based keypoints



(b) Edge based keypoints



(c) R based keypoints

Figure 5: Detected keypoints on different maps

Sziranyi, 2010a)) is in Figure 4. As R-function has lower values, the image can be better seen, if the natural logarithm is illustrated instead of the original map.

Results of the keypoint detection is in Figure 5, detected keypoints are in white. In case of intensity based and edge based difference map, keypoints are mainly located around newly built objects, but there are some false points. Intensity and edginess are too sensitive to illumination change, so altering contrast and color conditions result the appearance of false edges and corner points and the vanishing of real ones in the difference map.

In case of R (see Figure 5(c)), keypoint candidates covers nearly all buildings, and only a few points are in false areas.

After having some keypoint candidates indicating newly built objects, keypoints defining real changes should be selected somehow.

3 FILTRATION OF KEYPOINT CANDIDATES BASED ON LOCAL DESCRIPTOR

3.1 Detection of local structures

According to (Kovacs and Sziranyi, 2009), local contours around keypoint candidates provide adequate description of local characteristic at lower dimension than SIFT ((Lowe, 2004)), therefore they are efficient tool for matching and distinguishing. By generating local contours, we try to filter out false keypoints generated in the previous step. Surroundings of the candidate on the older and newer image (Figure 2(a) and 2(b)) are investigated and then compared.

The main steps for estimating local structure characteristics are the following:

1. Generating the Local Contour around keypoints for the original images (Xu and Prince, 1997)
2. Calculating Modified Fourier Descriptor (MFD) for the estimated closed curve (Rui et al., 1998)
3. Describing the contour by a limited set of Fourier coefficients (Licsar and Sziranyi, 2005)

As the specification shows, after detecting the keypoint candidates (the method is briefly summarized in Section 2.2), GVF Snake (Xu and Prince, 1997) was used for local contour (LC) analysis (Step 1). LC was computed in the original image, in a 20×20 size area, where the keypoint was in the middle. The generated LC assigns an individual shape to every keypoint, but the dimension is quite high. Therefore modified Fourier descriptors were applied (Step 2) with dimension reduction. This represents the LC at low dimension (Step 3). We have determined the cut-off frequency by maximizing the recognition accuracy and minimizing the noise of irregularities of the shape boundaries and chose the first twenty coefficients (excluding the DC component to remove the positional sensitivity). This reduction technique resulted a low dimensional descriptor (LDD), representing the local information around the keypoint candidate.

3.2 Finding similar local curves around keypoints

Our assumption was that after having the LDDs for the keypoints, differences between keypoint surroundings can be searched through this descriptor set. We extended the MFD method to get symmetric distance computation as it is written in (Kovacs and Sziranyi,



Figure 6: Result of Local Contour matching

2009). By comparing a keypoint (p_i) on the older image ($p_{i,old}$) and on the newer image ($p_{i,new}$), $S(p_{i,old}, p_{i,new})$ represents the similarity value. If the following criteria exists:

$$S(p_{i,old}, p_{i,new}) > \epsilon_3 \quad (3)$$

where $\epsilon_3 = 3$ is a tolerance value, than the keypoint is supposed to be a changed area.

3.3 New edge map

After testing the algorithm, we realized that active contours with original intensity based edge map, are sensible to changes. Even for similar contours, the method often generated false positive results. This meant that changeless places were declared as newly built objects.

Instead of the f edge map and E_{ext} external force of the original GVF snake (Xu and Prince, 1997) we used the Harris normalized R characteristic function (R_n) (Section 2.1). (The method is briefly described in (Kovacs and Sziranyi, 2010a).)

$$f_{R_n}(x, y) = G_\sigma(x, y) * R_n(x, y) \quad (4)$$

Detected contours are smoother and more robust in case of the R_n function. We benefit from this smoothness, as contours can be better distinguished. However, as there is no real contour in the neighbourhood of the keypoints, AC-method is only used for exploiting the local information to get low-dimensional descriptor, therefore significance of accuracy is overshadowed by efficiency of comparison. Remaining points after can be seen in Figure 6.

3.4 Enhancing the number of saliency points

After selecting the saliency points indicating changes, more featuring points are to be gained to enhance the number of keypoints around the detected changes. Therefore we are looking for saliency points that are not presented in the older image, but exist on the newer one. We apply the Harris corner detection method again with some modification (see (Kovacs and Sziranyi, 2010a)). This modification emphasizes both edge and corner points of the image.

By calculating saliency points for older and newer image as well, an arbitrary $q_i = (x_i, y_i)$ point is selected if it satisfies all of the following conditions:

- (1.) $q_i \in H_{new}$
- (2.) $q_i \notin H_{old}$



Figure 7: Enhanced number of Harris keypoints

$$(3.) d(q_i, p_j) < \epsilon_4$$

H_{new} and H_{old} are the sets of keypoints generated in the newer and older image, $d(q_i, p_j)$ is the Euclidean distance of q_i and p_j , where p_j denotes the point with smallest Euclidean-distance to q_i selected from H_{old} .

New points are searched iteratively, with $\epsilon_4 = 10$ condition. Here, ϵ_4 depends on the resolution of the image and on the size of buildings. If resolution is smaller, than ϵ_4 has to be chosen as a smaller value.

Figure 7 shows the enhanced number of keypoints.

3.5 Reconciling edge and corner detection

Now an enhanced set of saliency points is given, denoting the possible area of changes, which can be the basis of object detection. We redefine the problem in terms of graph theory. (Sirmacek and Unsalan, 2009)

A graph G is represented as $G = (V, E)$, where V is the vertex set, E is the edge network. In our case, V is already defined by the enhanced set of Harris points. Therefore, E needs to be formed.

Information about how to link the vertices can be controlled by classical edge maps. These maps can help us match only vertices belonging to the same objects.

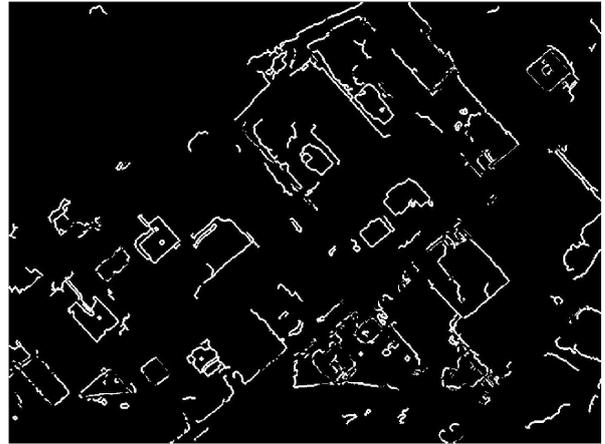
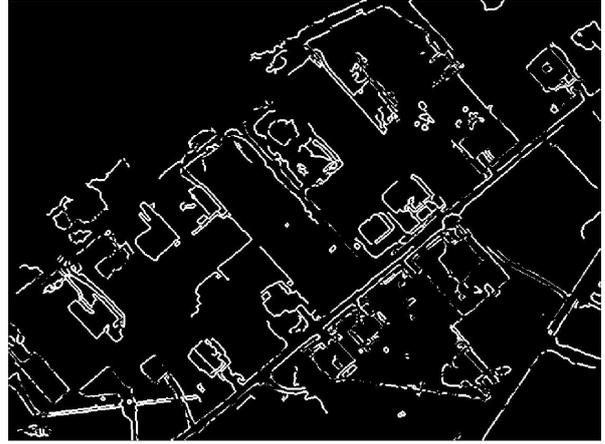
By generating the R and u^* components (mentioned in Subsection 2.2) of the original image pair, Canny edge detection (Canny, 1986) with large threshold ($Thr = 0.4$) is applied on them. The difference of these edge maps shows the edges appearing on the newer image.

C_r and C_u marks the edge maps generated by Canny detector. (Figure 8(a) and 8(b))

The process of matching is as follows. Given two vertices: $v_i = (x_i, y_i)$ and $v_j = (x_j, y_j)$. We match them if they satisfy the following conditions:

- (1.) $d(v_i, v_j) = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} < \epsilon_5$,
- (2.) $C_{\dots}(x_i, y_i) = \text{true}$,
- (3.) $C_{\dots}(x_j, y_j) = \text{true}$,
- (4.) \exists a finite path between v_i and v_j .

C_{\dots} indicates either C_r or C_u . ϵ_5 is a tolerance value, which depends on the resolution and average size of the objects. We apply $\epsilon_5 = 30$.

Figure 8: Detection of new edges: (a) on R component, (b) on u^* component.

These conditions guarantee that only vertices connected in the edge map are matched.

We obtain a graph composed of many separate subgraphs, which can be seen in Figure 9. Each of these connected subgraphs is supposed to represent a building. However, there might be some unmatched keypoints, indicating noise. To discard them, we select subgraphs having at least two vertices.

To determine the contour of the subgraph-represented buildings, we used the aforementioned GVF snake method. The convex hull of the vertices in the subgraphs is applied as the initial contour.

4 EXPERIMENTS AND CONCLUSIONS

Result of the contour detection can be seen in Figure 10 and 11.

The main advantage of our method is that it does not need any building or shape templates (like in (Benedek, 2010)) and can detect objects of any size and shape.

The algorithm also detects some keypoints of the newly made roads and routes (see Figure 7). This information could be used efficiently for road detection. Certainly, when detecting routes, active contour method should be substituted for other process (like Hough transformation).

The method has difficulties in finding objects with similar color to the background, for example, it often misses buildings with dark, greyish roof. Sometimes one object is described with more



Figure 9: Subgraphs given after matching procedure



(a) Original image

(b) Detected changes

Figure 10: Result of the contour detection

than one subgraphs. These problems need to be solved in a forthcoming semantic or object evaluation step.

ACKNOWLEDGEMENTS

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

REFERENCES

- Benedek, C., 2010. Efficient building change detection in sparsely populated areas using coupled marked point processes. In: IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Honolulu, Hawaii, USA.
- 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.



Figure 11: Result of the contour detection

Benedek, C. and Szirányi, T., 2009. Change detection in optical aerial images by a multi-layer conditional mixed markov model. *IEEE Trans. Geoscience and Remote Sensing* 47(10), pp. 3416–3430.

Canny, J., 1986. A computational approach to edge detection. *IEEE Trans. Pattern Analysis and Machine Intelligence* 8(6), pp. 679–698.

Ghosh, S., Bruzzone, L., Patra, S., Bovolo, F. and Ghosh, A., 2007. A context-sensitive technique for unsupervised change detection based on hopfield-type neural networks. *IEEE Trans. Geoscience and Remote Sensing* 45(3), pp. 778–789.

Harris, C. and Stephens, M., 1988. A combined corner and edge detector. In: *Proceedings of the 4th Alvey Vision Conference*, pp. 147–151.

Kovacs, A. and Sziranyi, T., 2009. Local contour descriptors around scale-invariant keypoints. In: *International Conference on Image Processing*, Cairo, Egypt, pp. 1105–1108.

Kovacs, A. and Sziranyi, T., 2010a. High definition feature map for gvf snake by using harris function. In: *Advanced Concepts for Intelligent Vision Systems*, Springer - Lecture Notes on Computer Sciences, Sydney, Australia.

Kovacs, A. and Sziranyi, T., 2010b. 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.

Lafarge, F., Descombes, X., Zerubia, J. and Pierrot Deseilligny, M., 2008. Automatic building extraction from dems using an object approach and application to the 3d-city modeling. *ISPRS Journal of Photogrammetry and Remote Sensing* 63(3), pp. 365–381.

Licsar, A. and Sziranyi, T., 2005. User-adaptive hand gesture recognition system with interactive training. *Image and Vision Computing* 23(12), pp. 1102–1114.

Lowe, D. G., 2004. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 60, pp. 91–110.

Peng, T., Jermyn, I. H., Prinnet, V. and Zerubia, J., 2008. Incorporating generic and specific prior knowledge in a multi-scale phase field model for road extraction from vhr images. *IEEE Trans. Geoscience and Remote Sensing* 1(2), pp. 139–146.

Rui, Y., She, A. and Huang, T., 1998. A modified fourier descriptor for shape matching in MARS. In: *Image Databases and Multimedia Search*, pp. 165–180.

Shah, C., Sheng, Y. and Smith, L., 2008. Automated image registration based on pseudoinvariant metrics of dynamic land-surface features. *IEEE Trans. Geoscience and Remote Sensing* 46(11), pp. 3908–3916.

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.

Wiemker, R., 1997. An iterative spectral-spatial bayesian labeling approach for unsupervised robust change detection on remotely sensed multispectral imagery. In: *Computer Analysis of Images and Patterns*, Springer - Lecture Notes on Computer Sciences, pp. 263–270.

Xu, C. and Prince, J. L., 1997. Gradient vector flow: A new external force for snakes. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 66–71.

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.