ROBUST CHANGE DETECTION BY INTEGRATING OBJECT-SPECIFIC FEATURES

Chunlei Huo[a, *], Keming Chen[a], Zhixin Zhou[b, a], Hanqing Lu[a], Jian Cheng[a]

[a] National Laboratory of Pattern Recognition, Institute of Automation, CAS, Beijing, 100190 – (clhuo, kmchen, luhq, jcheng)@nlpr.ia.ac.cn

[b] Beijing Institute of Remote Sensing, Beijing, 100854 – zhixin.zhou@ia.ac.cn

Commission VII, WG VII /5

KEY WORDS: Urban, Change Detection, High Resolution Images, Object-Specific Features, Feature Fusion, Quickbird.

ABSTRACT:

Urban change detection of high resolution images is very difficult due to the complex nature of the urban environments. A robust approach is proposed in this paper for urban change detection of high resolution images, which is based on the integration of object-specific features. To model the contextual information and improve the interclass variability between different classes, homogeneous regions are first extracted. Based on the robust object-specific difference measure, the complex changes of urban scene are represented by the fusion of object-specific spectral and textural features. Experimental results obtained on QuickBird images confirm the effectiveness and advantage of the proposed approach.

1. INTRODUCTION

High and accelerating rate of urban changes, in particular in the developing countries, calls for an efficient and fast technique for mapping the changes with the required accuracy for updating the existing topographic urban geospatial databases. With the development of HR (High Resolution) sensors, it is possible to detect changes at the smaller spatial scale. However, urban change detection of HR images is more difficult due to the complex nature of the urban environments. First, a wide range of materials (e.g. concrete, asphalt, metal, shingles, water, grass, trees and soil) can be found in the urban environment of HR data, which provides detailed information for a better characterization of the area to be considered. On the other hand, such materials create additional problems in terms of feature extraction. Second, due to the resulting high within-class spectral variance, the urban areas are often more heterogeneous than natural landscapes, so the traditional approaches (Coppin et al., 2004, Lu et al., 2004, Radke et al., 2005) cannot always provide the reliable result. It is the difficulties that make most of the change-detection processes implemented manually (such as on-screen change detection), which are time consuming and expert dependent.

To address the above difficulties, in this paper, we propose a robust approach to urban change detection of HR images. By virtue of object-specific features, the interclass variability between different classes is improved. By taking advantages of the combination of spectral and textural features, the statistical separability between the changed class and the unchanged class is increased. Moreover, due to the powerful ability provided by the robust difference measure, the proposed approach is discriminative to detect the structural changes and robust to the illumination variation.

The paper is organized in four sections. Section 2 presents a detailed description of the proposed approach step by step. Section 3 reports the experimental results obtained on real QuickBird images. Finally, section 4 draws the conclusions of this paper.

2. THE PROPOSED CHANGE DETECTION APPROACH

The rationale of the proposed approach is to extract homogeneous regions from images by segmentation, to represent complex objects by object-specific features and to achieve reliable results by effective and robust difference measure. A more detailed description of all the steps of the proposed approach is elaborated in the following paragraphs.

2.1 Object Extraction

One of the limitations of traditional pixel-based change detection approaches is the difficulty to model the contextual information for every pixel by the moving window with the size and shape particular to the corresponding object (Blaschke and Strobl, 2001). To address this problem, the first step of the proposed method is to divide the images into some homogeneous regions, which allows for extracting the contextual features without being affected by "edge effects". In this paper, we use segmentation techniques based on Mean shift (Comaniciu and Meer, 2002) to extract objects due to its good performance and the relative freedom from specifying an expected number of segments. Mean shift begins at each pixel and estimates the local density of similar pixels. More specifically, Mean shift algorithm estimates the local density gradient of similar pixels. These gradient estimates are used within an iterative procedure to find the peaks in the local density. All pixels that are drawn upwards to the same peak are then considered to be members of the same segment. The detailed algorithm about Mean shift can be found in (Comaniciu and Meer, 2002).

For two co-registered images X_1 and X_2, to avoid the

* Corresponding author.

"multiplication effect" of post-classification-based change detection and reduce the computational cost, we apply Mean shift to the image (e.g. X_1) which is better in quality, to generate the object-region set $R=\{R_1,\ldots,R_M\}$. By object extraction, the spectral variability between different classes is improved and object-specific features can be extracted to describe the complex objects.

2.2 Object-Specific Comparison

Local structure features, or texture, is very important to describe the complex man-made objects. Moreover, it is stable with respect to illumination changes and noise. However, the texture will become less valid than the simple spectral features for the homogeneous regions of urban environments. Therefore, it is desirable to integrate the texture difference and the spectral difference for robust change detection.

2.2.1 Spectral Difference: Let d_s be the spectral difference images, each $d_s(R_i)(1\le i\le M)$ of which is computed based on the spectral difference within the object R_i between two images. To be robust to noise, the spectral difference measure should be normalized into the range of [0, 1] before integration. This can be done by applying the following slope function

$$D_s(R_i)=\begin{cases} 1 & : \quad |d_s(R_i)|>2T \\ |d_s(R_i)|/(2T) & : \quad otherwise \end{cases} \tag{1}$$

The parameter T should be properly chosen to cope with image noise. Since the image noise can be modeled as Gaussian noise $N(0,\sigma)$, the noise in the difference image has a Gaussian distribution $N(0,\sigma_d)$ with $\sigma_d=\sqrt{2}\sigma$. However, due to the effect of illumination changes, the shifts of spectral values for unchanged regions should be compensated to compute σ_d. First, the shift of gray value at each region is calculated as

$$\overline{d}_s(R_i)=\frac{1}{N}\sum_{p\in R_i} d_s(p) \tag{2}$$

Where N is the number of pixels in the region R_i. Then, the difference image becomes

$$d_s{}'=d_s-\overline{d_s} \tag{3}$$

By estimating the variance of noise to image $d_s{}'$, σ_d can then be obtained. The average shift of spectral difference for unchanged regions is estimated as

$$\overline{d_u}=\frac{1}{\|\Omega\|}\sum_{R_i\in\Omega}\overline{d_s}(R_i) \tag{4}$$

With $\Omega=\{R_i : |d_s{}'(R_i)|<2\sigma_d$ and $|d_s(R_i)|<T_{med}\}$, where T_{med} is the median of $d_s(R_i)$ and $\|\Omega\|$ is the number of regions in the set Ω, most objects of which come from the unchanged regions. From these estimations, the parameter T can be chosen as $|\overline{d_u}|+3\sigma_d$.

2.2.2 Textural Difference: Texture is an important feature since it represents the spatial arrangement about gray levels of pixels in an object. Among different texture features, the gradient value is a good measure since it describes how the gray level changes within a neighborhood and is less sensitive to illumination changes. For this reason, in this paper, the gradient vector is used to measure the local texture difference.

Let $f_i(p)$ be the *i*th image and $f_i{}'(p)=[f_i{}^x(p) \quad f_i{}^y(p)]$ the gradient vector at pixel p, where $f_i{}^x(p)=\nabla_x f_i(p)$ and $f_i^y(p)=\nabla_y f_i(p)$. For an object region R_i, the following distance is used to measure the textural difference

$$d_t(R_i)=1-\frac{\sum_{p\in R_i}2C_{12}(p)}{\sum_{p\in R_i}(C_{11}(p)+C_{22}(p))}. \tag{5}$$

Where $C_{ij}(p)$ is the cross-correlation of gradient vectors of two images at p. Such difference measure can describe the difference within the local regions, both the difference in magnitude and the difference in angle. What's more, as illustrated in (Li and Leung, 2002), it is robust to illumination changes and noise. Different from the original work in (Li and Leung, 2002), we compute the difference in an object-specific fashion, the contextual information is considered without being perturbed by the "edge effects" caused by the fixed moving window. As illustrated in Experiment section, it is the object-specific fashion that helps improve the overall performance. However, if the object-region R_i in both images is homogeneous, the above difference measure becomes invalid since the denominator of the second term in equation (5) would be very small. For this reason, a validity weight, $w(R_i)$, is computed for gradient difference within each object. Let

$$g(R_i)=\max_{i=1,2}\sqrt{\frac{1}{\|R_i\|}\sum_{p\in R_i}C_{ii}(p)} \tag{6}$$

Then

$$w(R_i) = \begin{cases} 1 & : \quad |g(R_i)| > 2T_w \\ |g(R_i)|/(2T_w) & : \quad otherwise \end{cases} \tag{7}$$

Where $\| R_i \|$ means the number of pixels in R_i, T_w is a parameter based on image noise distribution. As the work in (Li and Leung, 2002), we set T_w as $5\sigma_d$. From the above analysis, the texture difference between two images can be defined as

$$D_t(R_i) = w(R_i) \times d_t(R_i). \tag{8}$$

2.2.3 Fusion of Spectral and Textural Differences: Spectral and textural differences are two different views to represent the difference between two images. Considering the complementation of the spectral and textural differences, the basic strategy is therefore to integrate the information from these two sources.

Due to the noise and illumination changes, the texture difference D_t is considered to be more reliable and robust than the simple spectral difference D_s. Hence, we should depend on D_s only if the corresponding region has no texture. Different fusion approaches can be used, in this paper, two simple yet effective fusion strategies based on the weight of texture evidence are investigated. The first strategy combines the two types of differences based on the adaptive weight, i.e.,

$$d_{st}(R_i) = w_s(R_i) \times D_s(R_i) + w_t(R_i) \times D_t(R_i). \tag{9}$$

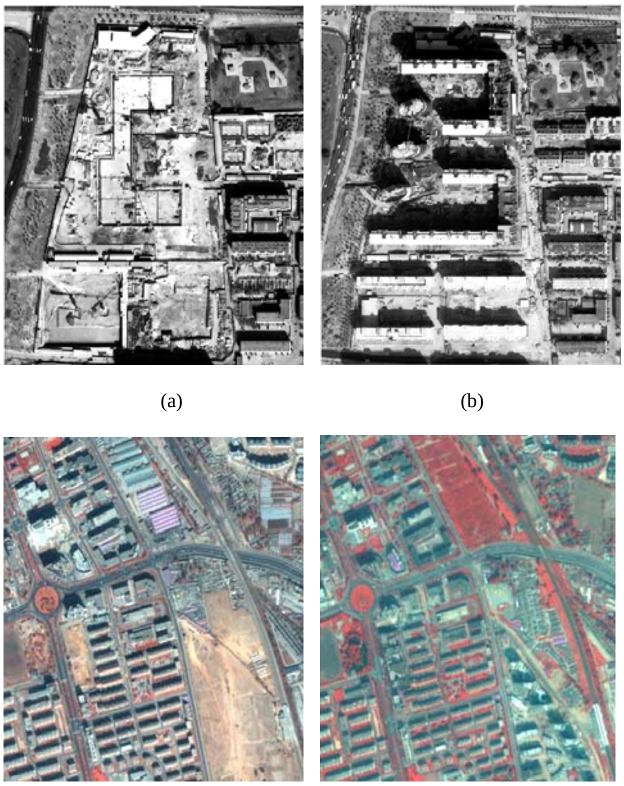
For simplicity, $w_t(R_i)$ can be set as $w(R_i)$, i.e., $w_t(R_i) = w(R_i)$ and $w_s(R_i) = 1 - w(R_i)$. While the second strategy combines the different change features based on the global weight, i.e.,

$$d_{st}(R_i) = (1-c) \times D_s(R_i) + c \times D_t(R_i). \tag{10}$$

Where c is a constant to denote the contribution of textural difference, it can be determined empirically. Finally, the changes can be detected by applying automatic threshold selection techniques (Rogerson, 2002) on d_{st}. The comparison of these two fusion strategies will be discussed in the next section.

3. EXPERIMENTAL RESULTS

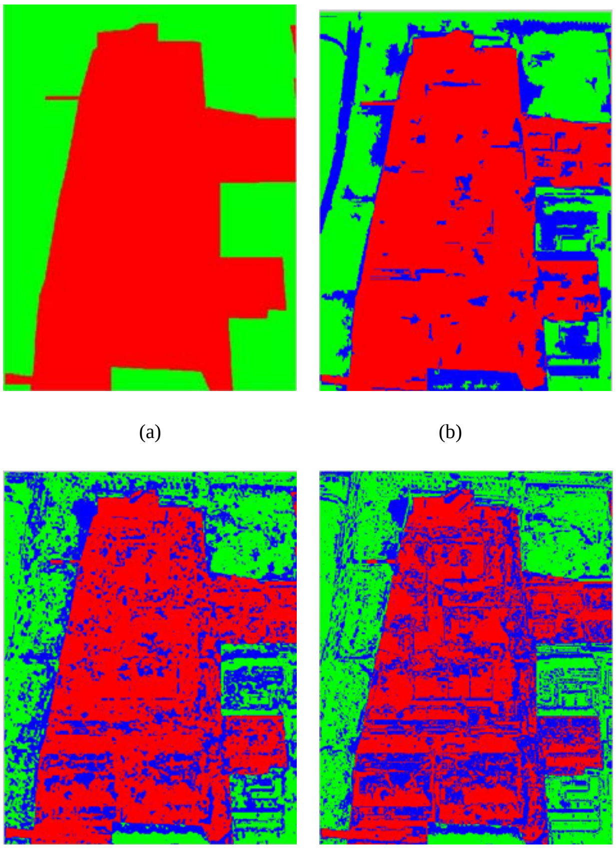
To assess the effectiveness of the proposed approach, experiments were conducted on HR images of many scenes. For space limitation, only the results on two data sets are shown in this paper. The first data set is composed of two panchromatic images of 700×500 pixels (61 cm/pixel), while the second data set is made up of two images of 500× 500 pixels (2.4 m/pixel) and composed of three channels (red, green and blue). Both data sets are taken over Beijing (China) acquired by QuickBird in 2002 and 2003. The selected test sites are shown in Fig. 1. Fig. 2(a) and Fig. 3(a) show the ground truth of the two data sets that are labeled manually, in which the changed class and the unchanged class are shown in red and green respectively. For comparison visually, in other three images of Fig. 2 and Fig. 3, we use the color of red, green and blue to represent the changed class detected correctly, the unchanged class detected correctly and the mis-classified class respectively. Fig. 2(b) and Fig. 3(b) are the results by the proposed approach based on the optimal threshold. To demonstrate the effectiveness of the proposed approach, the change detection results by other two different methods are also shown.



(a) (b) (c) (d)

Figure 1: Data sets used in this paper. (a) and (b): the first data set, (c) and (d): the second data set.

Figure. 2(c) and Fig. 3(c) give the results based on the combination of spectrum and texture difference measures but in a pixel fashion, while Fig. 2(d) and Fig. 3(d) show the results based on pixel-based comparison and only on spectrum difference. For convenience, we abbreviate the three above approaches as “object + spectrum + texture”, “pixel + spectrum + texture” and “pixel+ difference” respectively. By comparison of Fig. 2(b), (c), (d) and Fig. 2(a), we can see the advantages of the proposed approach over the other two procedures. First, for high resolution images, object-based change detection outperforms pixel-based change detection, which can be concluded by comparing Fig. 2(b) and (c). Although the same features are used to measure the difference, however, pixel-based change detection is less powerful to model the contextual information by the moving window of fixed shape and size, while object-based change detection enables to obtain the features specific to the objects without taking into account nearby regions. One the one hand, the statistical separability between the changed class and the unchanged class is improved by object-specific feature extraction and comparison. As a result, object-based change detection produces the result more similar to the reference change map. On the other hand, for pixel-based change detection, changed pixels are mixed up with the unchanged pixels, the linear discriminability between the changed class and the unchanged class is reduced and more change detection errors are involved. Moreover, many isolated pixels are labeled as changed or unchanged, which is difficult to interpret. Similar conclusions can be drawn from the comparison of results on the second data set. Furthermore, in this case, traditional pixel-based change detection based on spectrum difference is much worse, which shows the importance of texture difference measure to the complex changes of urban scene.

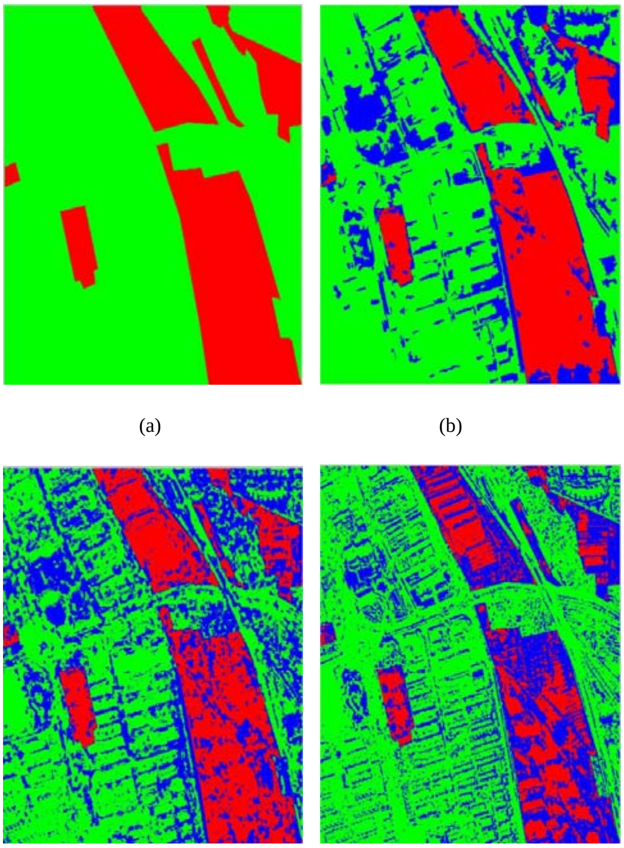


(a) (b)

(c) (d)

Figure 2: Results comparison on the first data set. (a): reference change map. (b): change map by object-based method with spectrum and texture differences. (c): change map by pixel-based method with spectrum and texture differences. (d): change map by pixel-based method with only spectrum difference.

In order to evaluate the quality of a change image independently of the choice of the thresholding algorithm, the evolution of the detection probability (Pdet) as a function of the false-alarm probability (Pfa) may be evaluated, where a set of constant thresholds is applied to the whole change image. These are the so-called Receiver Operating Characteristics (ROC), and the plots of Pdet (Pfa) are called the ROC plots (Inglada and Mercier, 2007). To evaluate the proposed approach in detail, two sets of experiments are designed. The first set of experiments was aimed at comparing the performances of different fusion strategies presented in this paper. In the second set of experiments, we compared the ROC plots of the proposed method with other related techniques.



(a) (b)

(c) (d)

Figure 3: Results comparison on the second data set. (a): reference change map. (b): change map by object-based method with spectrum and texture differences. (c): change map by pixel-based method with spectrum and texture differences. (d): change map by pixel-based method with only spectrum difference.

In ideal case, fusion strategy based on adaptive weight (9) outperforms that on global weight (10) since the former considers the evidence of texture for each object in combining measures. However, this conclusion holds only when the spectral difference of unchanged regions varies within a range. To further investigate the two fusion strategies, we compare the different performances of them on the second data set, in which spurious changes caused by seasonal changes are very difficult to remove. Fig. 4 shows the ROC plot, from which we can see that the performance of the global weight with c = 0.8 and c = 0.9 is more reliable than that of the adaptive weight. The underlying reason is that when the illumination between images exceeds the range, the shift of spectral values from unchanged regions may violate from Gaussian distribution, which further affects the parameter T_w. The result of c = 0.7 is less reliable than that of c = 0.8 since the importance of the texture is not given enough emphasis. Despite of the importance of texture, however, the difference measure on the texture purely (c = 1) cannot achieve the most reliable result, which demonstrates the importance of feature fusion.

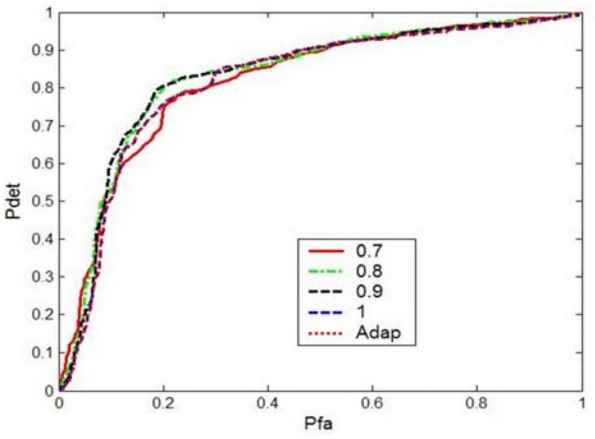


Figure 4: Performance comparison against different fusion strategies.

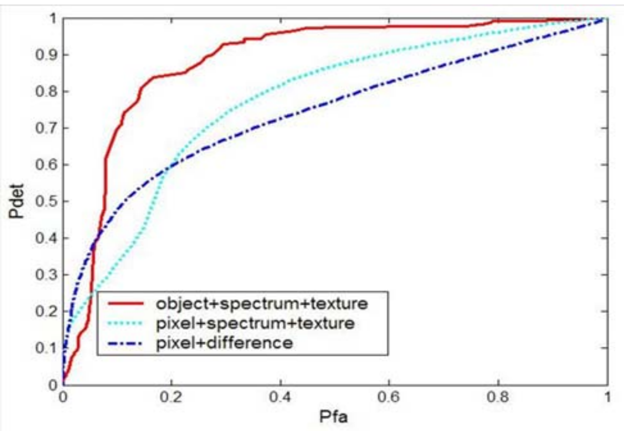


Figure 5: Performance comparison on the first data set against different methods.

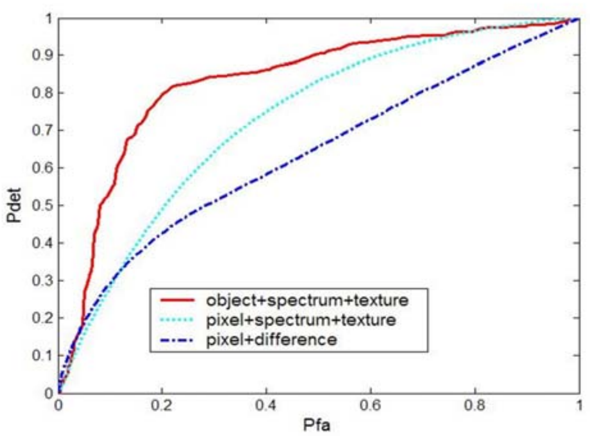


Figure 6: Performance comparison on the second data set against different methods.

Figure. 5 and Figure. 6 give the ROC plots using the ground truth shown in Fig. 2(a) and Fig. 3(a). Fig. 5 shows the ROC plots of the first data set based on the adaptive weight. While Fig. 6 gives the ROC plots of the second data set based on the global weight with c = 0.8. As can be seen from Fig. 5 and Fig.6, the proposed approach is more reliable than the other two approaches. This comparison indicates the effectiveness of our approach. Such effectiveness depends mainly on the powerful ability provided by the object-based change detection and robust feature extraction. In detail, object-based region extraction improves the statistical separability between different land-cover classes, while robust feature extraction and object-specific feature fusion can increases the statistical separability of the changed class from the unchanged class. Thanks to this ability, the proposed approach turns out to be superior to the traditional method.

4. CONCLUSIONS

Conventional pixel-based change detection schemes are less suited for urban change detection of HR images due to the resulting high intraclass and low interclass variability as well as complex man-made objects. In this paper, a robust approach is proposed for urban change detection of HR images, the main innovation lies in the incorporation of the object-based change detection and robust feature extraction as well as feature fusion. The experiment results reported in this paper confirm the effectiveness of the presented techniques. Despite of the promising preliminary results, many future developments need be considered to make our approach more robust and more perfect.

ACKNOWLEDGEMENTS

This work was supported by Natural Science Foundation of China (60121302, 60605004).

REFERENCES

Blaschke, T. and Strobl, J., 2001. What's wrong with pixels? some recent developments interfacing remote sensing and GIS. GIS Zeitschrift fur Geoinformationssysteme, 6, pp. 12–17.

Comaniciu, D. and Meer, P., 2002. Mean shift: A robust approach toward feature space analysis. IEEE Trans. on Pattern Analysis and Machine Intelligence, 24(5), pp. 603–619.

Coppin, P., Lambin, E. and Jonckheere, I., 2004. Digital change detection methods in natural ecosystem monitoring: a review. International Journal on Remote Sensing, 25(9), pp. 1565–1596.

Inglada, J. and Mercier, G., 2007. A new statistical similarity measure for change detection in multitemporal SAR images and its extension to multiscale change analysis. IEEE Trans. Geoscience and Remote Sensing, 45(5), pp. 1432–1445.

Li, L. and Leung, M., 2002. Integrating intensity and texture differences for robust change detection. IEEE Transaction on Image Processing, 11(2), pp. 105–112.

Lu, D., Mausel, P., Brondizio, E. and Moran, E., 2004. Change detection techniques. International Journal on Remote Sensing, 25(12), pp. 2365–2407.

Radke, R. J., Andra, S., O.Al-Kofahi and Roysam, B., 2005. Image detection algorithms: a systematic survey. IEEE Trans. Image Processing, 14(3), pp. 294–307.

Rogerson, P. A., 2002. Change detection thresholds for remotely sensed images. Journal of Geographical Systems, 4, pp. 85–97.