

On Automatic Threshold Selection in Regression Method for Change Detection in Remote Sensing Images

R. Dianat* and S. Kasaei*, *Senior Member, IEEE*

* Department of Computer Engineering, Sharif University of Technology, Tehran, Iran
dianat@ce.sharif.edu, skasaei@sharif.edu

Abstract- Automatic threshold selection in regression method of change detection in remote sensing images is investigated in this paper. We have reviewed the three well-known automatic threshold selection methods in the image processing field; namely, convex hull, Otsu, and ISODATA. We have then described the adaptive threshold selection concept. We have then shown that from these three methods, only the ISODATA can be successfully applied to regression method and have proposed an ISODATA-based adaptive threshold selection method. Experimental results show the efficiency of the proposed adaptive threshold selection method.

I. INTRODUCTION

Change detection is an important process in many remote sensing applications; such as monitoring urban development and assessing damages. In the simplest case, we are given two spatially co-registered images captured from the same scene but in different time instances (and probably by using different capturing sensors) and aim at detecting the changes occurred at that time interval in the scene. The most straightforward approach to change detection is image subtraction; in which the two images are simply subtracted from each other. Then, the pixels with difference values greater than a predetermined threshold are considered as changed pixels. The subtraction method might seem efficient but it suffers from severe shortcomings such as high sensitivity to illumination variation, sensor gain changes, and so forth. Therefore, a variety of change detection methods have been introduced to mitigate these shortcomings. These include *post classification comparison* (PCC), *direct multivariate classification* (DMDC), *principal component analysis* (PCA), *multivariate alteration detection* (MAD), and *image regression*. Interested readers can refer to [1-3] for more details. Among these methods, the regression method has some interesting properties such as

resistance to variation in illumination and change of sensor gain. As can be seen later, one of the challenging parts of the regression method is the threshold assignment. Since it is desirable to have an automatic change detection system, one needs an automatic threshold selection method. In this paper, we have investigated this subject. We first survey three such methods; namely, convex hull, ISODATA, and Otsu. According to our conducted experiments, the ISODATA method outperforms the other threshold selection methods for remote sensing applications. Also, we have described the adaptive threshold selection criteria. In the three mentioned threshold selection methods, only one threshold is calculated for the entire image. But, it is better to have several thresholds each for a specific region of the image. (This concept is called adaptive threshold selection.) As such, we have designed an ISODATA-based adaptive threshold selection method and have shown its performance in regression method.

The rest of the paper is organized as follow. In Section II, the related work is described. The regression method and its advantages are discussed and three conventional threshold selection methods and the adaptive threshold selection concept are described. In Section III, the performance comparison of the conventional threshold selection methods is given and our adaptive threshold selection method is introduced. Finally, the conclusion is drawn in Section IV.

II. RELATED WORK

In this section, the regression method and three conventional threshold section methods are described.

A. Regression Method

In the polynomial regression problem, we are given some points (x_i, y_i) to fit an n -degree polynomial $y=p(x)$ into them, using

$$p(x) = a_0 + a_1x^1 + a_2x^2 + \dots a_nx^n. \quad (1)$$

Let y_i represent the original values and $\hat{y}_i = p(x_i)$ denote the estimated values. The most popular regression estimator dates back to Gauss and Legendre and corresponds to minimize

$MSS_1 = \sum_{i=1}^m (y_i - \hat{y}_i)^2$ [4]. The *ordinary least square* (OLS) method given by

$$A = \arg \min \sum_{i=1}^m (y_i - \hat{y}_i)^2 = \arg \min \sum_{i=1}^m (y_i - a_0 + a_1x_i^1 + a_2x_i^2 + \dots a_nx_i^n)^2. \quad (2)$$

$$A^T = [a_0 \ a_1 \ a_2 \ \dots a_n]$$

If we derivate MSS_1 with respect to a_k and set it to zero, after some manipulations we get

$$\begin{bmatrix} \sum_{i=1}^m 1 & \sum_{i=1}^m x_i & \sum_{i=1}^m x_i^2 & \dots & \sum_{i=1}^m x_i^n \\ \sum_{i=1}^m x_i & \sum_{i=1}^m x_i^2 & \sum_{i=1}^m x_i^3 & \dots & \sum_{i=1}^m x_i^{n+1} \\ \sum_{i=1}^m x_i^2 & \sum_{i=1}^m x_i^3 & \sum_{i=1}^m x_i^4 & \dots & \sum_{i=1}^m x_i^{n+2} \\ \dots & \dots & \dots & \dots & \dots \\ \sum_{i=1}^m x_i^n & \sum_{i=1}^m x_i^{n+1} & \sum_{i=1}^m x_i^{n+1} & \dots & \sum_{i=1}^m x_i^{2n} \end{bmatrix} * \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \dots \\ a_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m y_i \\ \sum_{i=1}^m x_i y_i \\ \sum_{i=1}^m x_i^2 y_i \\ \dots \\ \sum_{i=1}^m x_i^n y_i \end{bmatrix} \Rightarrow X * A = Y \Rightarrow A = X^{-1}Y \quad (3)$$

In change detection using the regression method, it is assumed that the second image (Y) has a polynomial relation with the first image (X), where the pixel values constitute the (x_i, y_i) points. We apply the following steps to determine the changed points:

1. Compute regression coefficients (*i.e.*, vector A).
2. Obtain \hat{y}_i , for each pixel i .

3. Calculate the difference $diff_i = y_i - \hat{y}_i$ (the regression error).

4. If $diff_i$ is greater than a threshold, the i^{th} pixel is considered as a changed point.

Now, we take our attention on the ideas behind the method. First, suppose that the two images (X and Y) are the same and have no differences. In other words, $Y=X$. In this situation, by using a polynomial regression with any degree (even with degree of 1) we conclude that $diff=0$ (*i.e.*, the regression method correctly detects no changes for all pixels).

In the next step, suppose that the scene has no change other than illumination condition change (*e.g.*, one image is taken at night and the other one at noon). In most cases, we can model this kind of change by a linear or polynomial model. For example, the day and night illumination change can be modeled by adding (subtracting) a constant value to (from) the image. For this case, if we use the regression method we will have $diff=0$ (*i.e.*, the regression method again leads to correct results). Note that the simple image differencing method might be completely misleading in this situation.

Finally, suppose that some actual changes have occurred in the scene while some changes in illumination condition and sensor gain have also encountered. For now, suppose that the amount of change in the scene is small. With this assumption, we can obtain the regression model for illumination condition or sensor gain. In fact, we can consider the point with high regression error (*i.e.*, $diff$) as a changed point. This is because these points are very apart from the regression model.

B. Conventional Threshold Selection Methods

In this subsection, we describe three well-known threshold selection methods; namely, convex hull, ISODATA and Otsu. Also, we describe the role of adaptive threshold selection on improving the performance of these methods.

Convex Hull Method

The convex hull of a curve is the smallest region that encloses the curve in such a way that one can connect any two arbitrary points of it with a direct line so that it completely lies in that region [5]. The convex hull threshold selection method was first introduced by Rosenfeld in [6]. Interested readers can refer to [7] to overview the different variations of this method. One of these is explained next.

A scheme for the method is depicted in Fig.1, in which the image histogram is considered to have two dominant peaks; one corresponds to background and the other to the object. The method consists of the following steps. Here, we have assumed that non-zero histogram elements start with a and end with b .

- Calculate the two peaks (M, N).

- Connect the points $(M, H(M))$ and $(N, H(N))$ to get line L .

- Threshold is a point between $[M, N]$ that has a maximum distance from line L .

For more details (e.g., the method for calculating the peaks) refer to [8].

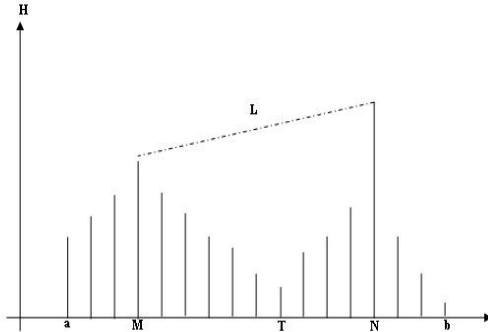


Fig. 1. Principle of convex hull method.

ISODATA Method

This iterative technique for threshold choosing was developed by Ridler and Calvard in [9]. The steps of the method are as follows:

- Select an initial threshold value T_0 (such as the half of maximum dynamic range).
- Segment the histogram into two parts so that one segment corresponds to foreground and the other to background.
- Compute the sample mean of gray values of foreground and background pixels ($m_{f,0}$ and $m_{b,0}$).
- Compute a new threshold value T_1 , as the average of these two sample means.
- Re-segment the histogram into two parts.
- Iterate the process, based upon the new threshold, until the threshold value converges. In other words

$$T_k = (m_{f,k-1} + m_{b,k-1}) / 2 \quad \text{until } T_k = T_{k-1} . \quad (4)$$

Otsu Method

This method was proposed by Nobuyuki Otsu in [10]. The idea of the method is to set the threshold so that the elements in a cluster locate near to each other as much as possible. As such, he defined the *within-class variance* as

$$\begin{aligned} \sigma_{\text{within class}}(T) &= \\ n_B(T)\sigma_B^2(T) + n_F(T)\sigma_F^2(T) & \quad (5) \\ n_B(T) &= \sum_{i=0}^{T-1} H(i) \quad n_F(T) = \sum_{i=T}^{255} H(i) \end{aligned}$$

in which σ_B and σ_F denote the variance of background and foreground elements, respectively. H is the histogram (or the probability density function). And, n_B and n_F are the probabilities of background and foreground clusters, respectively. The threshold is chosen so that $\sigma_{\text{within class}}$ is minimized.

As the minimization of this parameter is a difficult task, another criterion (named *between-class variance*) was defined by

$$\sigma_{\text{between classes}}^2(T) = \sigma^2 - \sigma_{\text{within class}}^2(T) \quad (6)$$

where σ^2 is the variance of all image elements, which is a constant (and not related to T). Consequently, for minimizing the within-class variance the between-class variance must be maximized. It can be proven that this variance is equal to

$$\begin{aligned} \sigma_{\text{between classes}}^2(T) &= \\ n_F(T)n_B(T)[\mu_F(T) - \mu_B(T)]^2 & \quad (7) \end{aligned}$$

The calculation of this relation is an easy and straightforward task. Thereupon, one should select T such that this relation gets maximized.

Adaptive Threshold Selection Method

Whereas the conventional threshold selection methods use a global threshold for all image pixels, the adaptive threshold selection changes the threshold dynamically over the image. In these approaches, typically, an image is divided into some non-overlapping blocks. The threshold for a specific pixel is obtained based on statistics of the block containing that pixel. In a general form, we may consider other neighboring blocks in addition. There are various types of adaptive threshold selection methods which use different methodologies for selecting the parameters and the type of adaptation (interested readers can refer to [7] for more details).

II. COMPARISON OF CONVENTIONAL METHODS AND INTRODUCING ADAPTIVE ISODATA-BASED METHOD

In Section II, we reviewed the three automatic threshold selection methods as well as the adaptive threshold concept. In this section, we first investigate the three mentioned methods in more detail and then introduce our proposed adaptive approach.

A. Drawbacks of Convex Hull and Otsu Methods

Drawback of convex hull method:

As stated in Section II, the main assumption in the convex hull method is that its histogram must have two dominant peaks. This results in an important difficulty on using it in the regression-based change detection processes. To better demonstrate this issue, the histogram of a typical regression difference image is depicted in Fig.2. As this figure shows, the histogram does not have two dominant peaks. Rather, this has a semi-Gaussian form. In fact, in the references related to regression methods, the regression error is always assumed to be Gaussian (e.g., refer to [4]). In this situation, the convex hull method cannot find the dominant peaks and will choose an improper threshold. In Table 1, we have listed the results of threshold selection by the three methods. As clearly can be seen, the convex hull has obtained a threshold which is improper and far away from the other two.

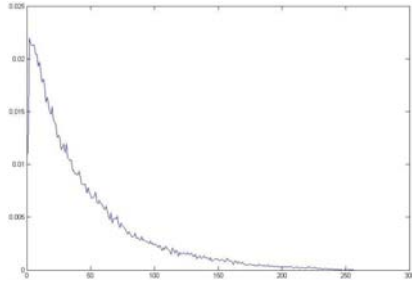


Fig. 2. Histogram of regression error for a typical regression difference.

TABLE I

COMPUTED THRESHOLD BY DIFFERENT THRESHOLD SELECTION METHODS.

Method	Computed Threshold
Convex Hull	247
ISODATA	67
Otsu	69

Drawback of Otsu method:

In Section II, we described the Otsu method. By considering (5)-(7), we found that in Otsu method we should calculate some statistics about the two classes. But, if in one class there are very small number of elements the method cannot successfully estimate these statistics and fails to obtain a proper threshold. In change detection literature, if the number of changes in the scene is small, the performance of Otsu method will be reduced. For better illustration we have performed the following experiment.

Experiment 1: In this experiment, we used two images from Bam city in Iran, taken in October 2002 and January 2004, respectively, before and after a terrible earthquake with degree 6.6 Richter. The size of images is 128×128 . In the scene, there are a small number of changes only in urban regions. We implemented the regression method and used the three threshold selection methods. The results are

shown in Fig. 3. These results clearly show the inefficiency of Otsu and convex hull methods.

From this experiment one can conclude that the convex hull and Otsu methods are not appropriate for change detection processes while in all of our implementations ISODATA method achieved reasonable results. Consequently, we adopted it as a basis of our proposed adaptive threshold selection approach.

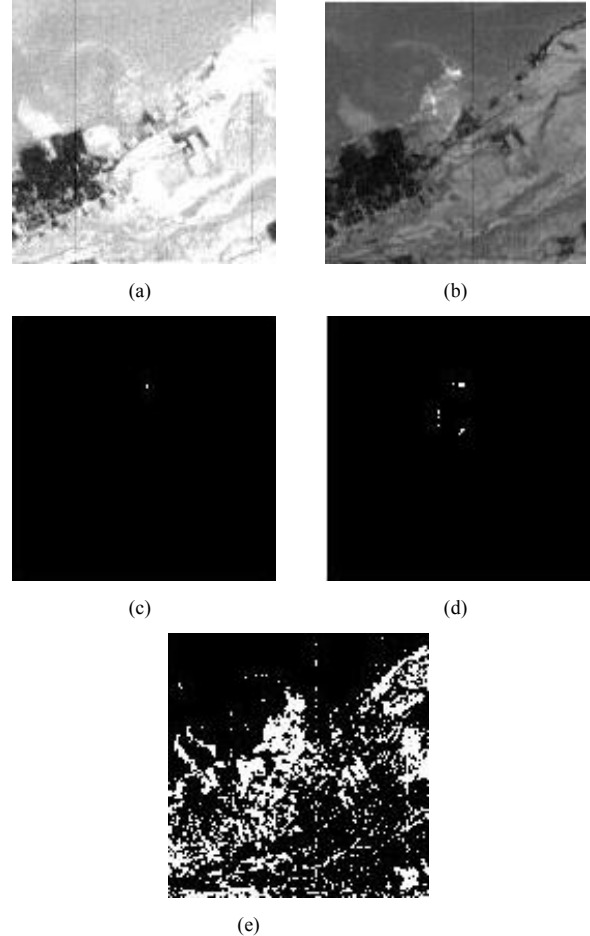


Fig. 3. Results of three automatic threshold selection methods: (a) - (b) original images, (c) convex hull, (d) Otsu, (e) ISODATA.

B. Proposed Adaptive Threshold Selection Method

Whereas the conventional threshold selection methods use a global threshold for all image pixels, an adaptive threshold selection method changes the threshold dynamically over the image. In these approaches, typically, an image is divided into some non-overlapping blocks. The threshold for a specific pixel is obtained based on the statistics of the block that contains it. There are various types of adaptive threshold selection approaches which use different methodologies for selecting the required parameters

(such as statistics and type of adaptation). For more details see [7].

The proposed adaptive threshold selection method is applied as follows.

- Divide the image into some non-overlapping blocks. (In our experiment we use 8×8 blocks.)
- For each block, obtain a threshold by means of a conventional method.
- Assign weights to each block. The weight is in inverse proportion to the distance of the block center from the pixel. In more exact terms, calculate the weight for block i by

$$w_i = \left(1 - \frac{d_i}{\sum_{j=1}^m d_j} \right) / (m-1) \quad i = 1, 2, \dots, m \quad (8)$$

where d_i indicates the distance of the center of block i from the pixel and m is the number of adjacent blocks plus the block that contains the pixel. In an 8-neighborhood system, this is at most 9 (because for some pixels, *e.g.* those lying on the border blocks, the adjacent blocks are less than 9). The idea behind such a weight is that the sum of all m weights must be equal to 1.

- For a specific pixel, the final threshold is obtained from

$$T = \sum_{i=1}^m w_i T_i. \quad (9)$$

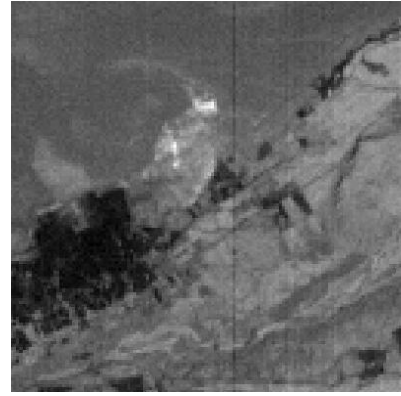
In this equation, m is the number of neighboring blocks. T_i and w_i are the threshold and the weight assigned to the i^{th} block, respectively.

We carried out an experiment to evaluate the performance of our method.

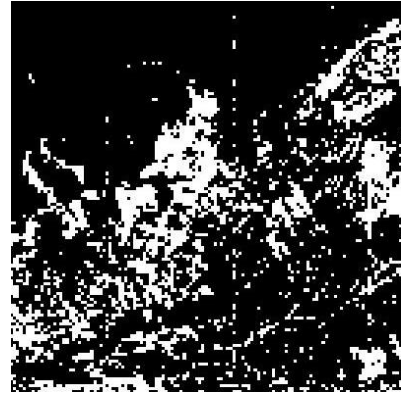
Experiment 2: In this experiment, we again used the Bam city images (for better viewing, the images are zoomed in). In Fig. 4, the results of the global ISODATA and our adaptive ISODATA are depicted. One can see that our method outperforms the global ISODATA method. For example, in regions labeled 1 and 2 that there is no change, our method has detected them correctly while the global ISODATA method has failed to detect them.



(a)

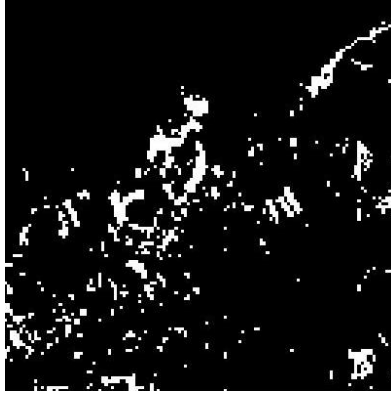


(b)



(c)

Fig. 4. Results of global ISODATA and proposed adaptive ISODATA threshold selection methods: (a) and (b) original images, (c) results of global ISODATA, (d) results of proposed adaptive ISODATA-based method.



(d)

Fig. 4. (continued).

III. CONCLUSION

In this paper, we investigated the automatic threshold selection process in regression-based change detection of remotely sensed images. At first, we described three conventional automatic threshold selection methods; namely, convex hull, Otsu and ISODATA. Then, we showed the shortcomings of the convex hull and Otsu methods which makes them improper to be used for remote sensing applications. Then, we described the adaptive threshold selection strategy and proposed an adaptive threshold selection method based on ISODATA method and showed its efficiency.

ACKNOWLEDGEMENT

This work was in part supported by a grant from ITRC.

REFERENCE

- [1] P. Coppin, I. Jonckheere, K. Nackaerts, and B. Muys, "Digital Change Detection Methods in Ecosystem Monitoring: a Review", *International Journal of Remote Sensing*, Vol. 25, No. 2, pp. 1565-1596, May 2004.
- [2] A. Nielsen, K. Conradsen, J. Simpson, "Multivariate Alteration Detection (MAD) and MAF Postprocessing in Multispectral, Bitemporal Image Data: New Approaches to Change Detection Studies", *Remote Sensing of Environment*, Vol. 64, pp. 1-19, 1998.
- [3] A. Nielsen, "The regularized Iteratively Reweighted MAD Method for Change Detection in Multi- and Hyperspectral Data", *IEEE Transactions on Image Processing*, Vol. 16, No. 2, pp. 463-478, February 2007.
- [4] P. Rousseeuw, "Robust Regression and Outlier Detection", John Wiley & Sons, 1987.
- [5] R.C. Gonzalez, R.E. Woods, "Digital image processing", Prentice Hall, 2002
- [6] A. Rosenfeld, P. De La Torre, "Histogram concavity analysis as an aid in threshold selection", *IEEE Transactions on Systems, Man and Cybernetics SMC-13*, pp. 231-235, 1983.
- [7] M. Sezgin, B. Sankur, "Survey over Image Thresholding Techniques and Quantitative Performance Evaluation", *Journal of Electronic Imaging*, Vol. 13, No. 1, pp. 146-168, 2004.
- [8] C.K. Lee, F.W. Choy, H.C. Lam, "Real-time Thresholding Using Histogram Concavity", *Industrial Electronics*, 1992, Proceedings of the IEEE International Symposium, pp. 500-503, May 1992.
- [9] T.W. Ridler, S. Calvard, "Picture thresholding using an iterative selection method", *IEEE Transactions on Systems, Man and Cybernetics SMC-8*, pp. 630-632, 1978.
- [10] N. Otsu, "A threshold selection method from gray level histograms", *IEEE Transactions on Systems, Man and Cybernetics SMC-9*, pp. 62-66, 1979.