Change-Detection Using Contextual Information and Fuzzy Entropy Principle

Maria Luiza F. Velloso, Flávio J. de Souza

Department of Electronics Engineering, Department of Computer Engineering
Uerj –Rio de Janeiro State University
Rua São Francisco Xavier, 524
20550-013, Rio de Janeiro,
BRAZIL
Phone: +55 21 2587 7442
mlfv@centroin.com.br, fjsouza@eng.uerj.br,

Abstract. This paper presents an unsupervised change detection method for computing the amount of changes that have occurred within an area by using remotely sensed technologies and fuzzy modeling. The discussion concentrates on the formulation of a standard procedure that, using the concept of fuzzy sets and fuzzy logic, can define the likelihood of changes detected from remotely sensed data. The fuzzy visualization of areas undergoing changes can be incorporated into a decision support system for prioritization of areas requiring environmental monitoring. One of the main problems related to unsupervised change detection methods lies in the lack of efficient automatic techniques for discriminating between changed and unchanged pixels in the difference image. Such discrimination is usually performed by using empirical strategies or manual trial-and-error procedures, which affect both, the accuracy and the reliability of the change-detection process. To overcome such drawbacks, in this paper, we propose an automatic technique for the analysis of the difference image. Such technique allows the automatic selection of the decision threshold. We used a thresholding approach by performing fuzzy partition on a two-dimensional (2-D) histogram, which included contextual information, based on fuzzy relation and maximum fuzzy entropy principle. Experimental results confirm the effectiveness of proposed technique.

1 Introduction

There has been a growing interest in the development of automatic change-detection techniques for the analysis of multitemporal remote sensing images. In the literature, two main approaches to the change-detection problem have been proposed: the supervised and the unsupervised. The former is based on supervised classification methods, which require the availability of a multitemporal ground truth in order to derive a suitable training set for the learning process of the classifiers. Although this approach exhibits some advantages over the unsupervised one, the generation of an appropriate multitemporal ground truth is usually a difficult and expensive task. Consequently, the use of effective unsupervised change-detection methods is fundamental in many applications in which a ground truth is not available. The unsupervised system approach is attractive for classifications tasks due to its self-organizing, generalizable, and fault-tolerant characteristics. In contrast to the supervised systems, the unsupervised system does not rely on user-defined training data and it is a advantageous characteristic because, frequently, there are not specialists in remote sensing or geoprocessing in the staff of city councils near sites that require environmental monitoring. Traditional methods of change detection using either air- or satellite-borne remotely sensed data also can be broadly divided in two categories: pre-classification and post-classification. Jensen [1] states that post-classification comparison of changes is the most commonly used method for quantitative analysis. It requires a complete classification of the individual dates of remotely sensed data, whereupon the operator produces a matrix of change that identifies 'from–to' land cover change classes. The main drawback with this method is errors in the individual data classification map will also be present in the final change detection. On the other hand, pre-classification methods detect changes due to variations in the brightness values of the images being compared. In any of the pre-classification approaches, the critical step relies on selecting appropriate threshold values in the lower and upper tails of the histogram representing values of change. This is so that areas of change can be accurately separated from those where no changes have occurred within the period of time considered. In all studies that create a change image, the value at which the threshold is set is somewhat arbitrary. In this paper, we work on one of the unsupervised change-detection techniques so-called "difference image". These techniques process the two multispectral images acquired at two different dates in order to generate a further image - the difference image. The values of the pixels associated with land cover changes present values significantly different from those of the pixels associated with unchanged areas. Changes are then identified by analyzing the difference image. In the widely used change vector analysis (CVA) technique [2], [3], [4], several spectral channels are used and, for each pair of corresponding pixels "spectral change vector" is computed as the difference between the feature vectors at the two times. Then, the pixel values in the difference image are associated with the modules of the spectral change vectors. So, the unchanged pixels present small gray-level values, whereas changed pixels present rather large values. In spite of their simplicity and widespread use, the described above change-detection methods exhibit a major drawback: a lack of automatic and nonheuristic techniques for the analysis of the difference image. An intuitive approach is to apply a grayscale threshold on the difference image – assume that the pixel values of the changed pixels are generally higher than the values of the unchanged pixels. If the histogram of the difference image is bimodal showing a peak for unchanged pixels and a peak for changed pixels, the appropriate value for the threshold can be either manually selected or statistically determined. However, due to the large variability on the change types and noise on the images, segmentation based on a single threshold usually performs poorly. Many methods have been proposed to select the thresholds automatically [5]. Most bilevel thresholding techniques can be extended to the case of multithresholding, therefore, we focus on a bilevel thresholding technique in this paper. The proposed approach will automatically determine the fuzzy region and find the thresholds based on the maximum fuzzy entropy principle. It involves a fuzzy partition on a two-dimensional (2-D) histogram where a 2-D fuzzy entropy is defined as in [6].

2 The 2-D Histogram

Let us consider two multispectral images, X1 e X2 acquired in the same geographical area at two different times, t1 e t2. Let us assume that such images have been coregistered. Let X represents the values of the pixels in the difference image obtained by applying the CVA technique to X1 and X2. For the sake of simplicity, the proposed technique will be presented in the context of the CVA method. However, a generalization to other methods based on the difference image is straightforward. In order to obtain a 2-D histogram of the difference image, we define the local average of a pixel, f(x; y), as the average intensity of its four neighbors denoted by g(x; y):

$$g(x,y)=\frac{1}{4}\left[f(x,y+1)+f(x,y-1)+f(x+1,y)+f(x-1,y)\right]+0.5. \quad (1)$$

A 2-D histogram can be viewed as a full Cartesian product of two sets X and Y, where X represents the gray levels of the difference image, f(x,y), and Y represents the local average gray levels, g(x,y): X=Y={0,1,2,…,L-1}, where L-1 is the higher level of intensity. This 2-D histogram is an array (L X L) with the entries representing the number of occurrences of the pair (f(x; y); g(x; y)). The pixels having the same intensity but different spatial features can be distinguished in the second dimension (local average gray level) of the 2-D histogram.
This paper is organized as follows. The next section presents the procedure used to generate the 2-D histogram. Section 3 describes the thresholding method used for segmentation in this work. Section 4 describes an evaluation experiment and discusses the results.

3 The Thresholding Method

Four fuzzy sets, ChangedX, NotChangedX, ChangedY, and NotChangedY, were defined based on the S-function and the corresponding Z-function as follows:

$$\begin{aligned}
ChangedX &= \sum_{x\in X}\frac{\mu ChangedX(x)}{x}=\sum_{x\in X}\frac{S(x,a,b)}{x}\\
ChangedY &= \sum_{y\in Y}\frac{\mu ChangedY(y)}{y}=\sum_{y\in Y}\frac{S(y,a,b)}{y}\\
NotChangedX &= \sum_{x\in X}\frac{\mu NotChangedX(x)}{x}=\sum_{x\in X}\frac{Z(x,a,b)}{x}\\
NotChangedY &= \sum_{y\in Y}\frac{\mu NotChangedY(y)}{y}=\sum_{y\in Y}\frac{Z(y,a,b)}{y}
\end{aligned} \quad (2)$$

where the S-function, is a S–shaped membership function, and is defined as

$$S(x,a,b)=\begin{cases}0 & \text{if } x\le a\\ 2\left(\dfrac{x-a}{b-a}\right)^2 & \text{if } a<x\le (a+b)/2\\ 1-2\left(\dfrac{b-x}{b-a}\right)^2 & \text{if } (a+b)/2<x\le b\\ 1 & \text{if } x\ge b\end{cases}, \tag{3}$$

and Z(x)=1 – S(x). The parameters a and b locate the extremes of the sloped portion of the curve, and Figure 1 shows a plot of a S-function with a=1 and b=8.

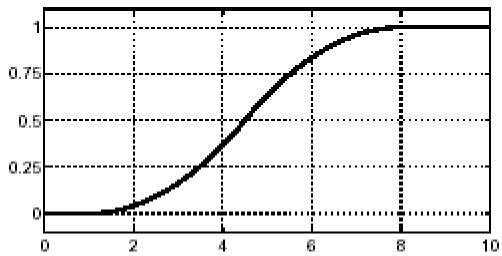


Fig. 1. A plot of a S-function with parameters a=1 and b=8.

The fuzzy relation Changed is a subset of the full Cartesian product space $X\times Y$, i.e., Changed = ChangedX × ChangedY $\subset X\times Y$

$$\mu Changed(x,y)=min(\mu ChangedX(x),\mu ChangedY(y)). \tag{4}$$

Similarly, NotChanged = DarkX × DarkY $\subset X\times Y$

$$\mu NotChanged(x,y)=min(\mu NotChangedX(x),\mu NotChangedY(y)). \tag{5}$$

Let A be a fuzzy set with membership function $\mu A(x_i)$, where $x_i, i=1,...,N$, are the possible outputs from source A with the probability $P(x_i)$. The fuzzy entropy of set A is defined as [7]

$$H_{fuzzy}(A)=-\sum_{i=1}^{N}\mu A(x_i)P(x_i)log(P(x_i)) \tag{6}$$

The image is divided in two blocks, NotChangedBlock, with $\mu NotChanged(x,y)>0$, and ChangedBlock, with $\mu Changed(x,y)>0$. The total image entropy is defined as

$$H(image)=H(NotChanged)+H(Changed). \tag{7}$$

The not changed block can be divided into a nonfuzzy region R_{NC} and a fuzzy region R_{NCF}.

$$NotChangedBlock = R_{NC} \cup R_{NCF}.$$

$$R_{NC} = \left\{(x,y) \middle| \mu NotChanged(x,y) = 1, (x,y) \in NotChangedBlock\right\},$$

$$R_{NCF} = \left\{(x,y) \middle| \; 0 < \mu NotChanged(x,y) < 1, (x,y) \in NotChangedBlock\right\}.$$

Similarly, the changed block is composed of a nonfuzzy region R_C and a fuzzy region R_{CF},

$$ChangedBlock = R_C \cup R_{CF}.$$

$$R_C = \left\{(x,y) \middle| \mu Changed(x,y) = 1, (x,y) \in ChangedBlock\right\},$$

$$R_{CF} = \left\{(x,y) \middle| \; 0 < \mu Changed(x,y) < 1, (x,y) \in ChangedBlock\right\}.$$

The following four entropies can be computed:

$$H(R_C) = -\sum_{(x,y)\in R_{CF}} \frac{\eta_{xy}}{\sum_{(x,y)\in R_C} \eta_{xy}}.log\left(\frac{\eta_{xy}}{\sum_{(x,y)\in R_C} \eta_{xy}}\right) \tag{8}$$

$$H(R_{CF}) = -\sum_{(x,y)\in R_{CF}} \mu Changed(x,y).\frac{\eta_{xy}}{\sum_{(x,y)\in R_{CF}} \eta_{xy}}.log\left(\frac{\eta_{xy}}{\sum_{(x,y)\in R_{CF}} \eta_{xy}}\right) \tag{9}$$

$$H(R_{NCF}) = -\sum_{(x,y)\in R_{NCF}} \mu NotChanged(x,y).\frac{\eta_{xy}}{\sum_{(x,y)\in R_{NCF}} \eta_{xy}}.log\left(\frac{\eta_{xy}}{\sum_{(x,y)\in R_{NCF}} \eta_{xy}}\right) \tag{10}$$

$$H(R_{NC}) = -\sum_{(x,y)\in R_{NC}} \frac{\eta_{xy}}{\sum_{(x,y)\in R_{NC}} \eta_{xy}}.log\left(\frac{\eta_{xy}}{\sum_{(x,y)\in R_{NC}} \eta_{xy}}\right) \tag{11}$$

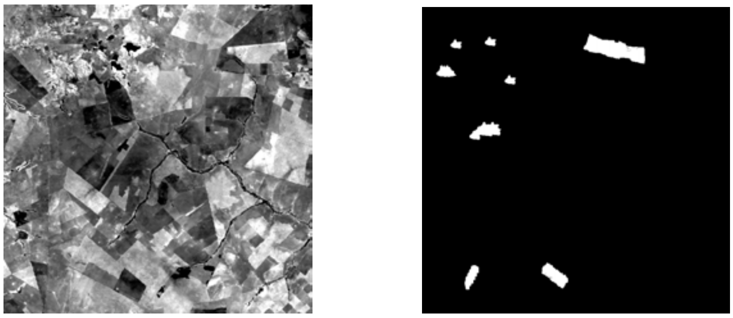
where η_{xy} is the element in the 2-D histogram which represents the number of occurrences of the pair (x; y). The membership functions $\mu Changed(x,y)$ and $\mu NotChanged(x,y)$ are defined in (4) and (5), respectively. It should be noticed that

the probability computations $\eta_{xy} / \sum \eta_{xy}$ in the four regions are independent of each other. To find the best set of a and b is an optimization problem, which can be solved by: heuristic searching, simulated annealing, genetic algorithm, etc. The proposed method consists of the following four major steps:
1) make the difference image;
2) find the 2-D histogram of the difference image;
3) Perform fuzzy partition on the 2-D histogram;
4) Compute the fuzzy entropy.
Steps 1) and 2) needs to be executed only once while steps 3) and 4) are performed iteratively for each set of (a; b). The optimum (a; b) determines the fuzzy region (i.e., interval [a; b]). The threshold is selected as the crossover point of the membership function which has membership 0.5 implying the largest fuzziness. Once the threshold vector (s; t) is obtained, it divides the 2-D histogram into four blocks, i.e., a Not Changed block, Block 0, a Changed block, Block1, and two noise (edge) blocks, Block2 and Block3, as shown in Figure 2. The bright extraction method is expressed as [6]

$$f_{s,t}(x, y, changed) = \begin{cases} 1 & f(x, y) > t \wedge g(x, y) > s \\ 0 & otherwise. \end{cases}$$

4 Experimental Results

In order to evaluate the robustness of the proposed technique for the analysis of the difference image, we considered a synthetic data set artificially generated. An image acquired by the Landsat-7 Thematic Mapper (TM) sensor, composed of bands 3, 4 and 5, in the middle west of Brazil was used as the reference image. In particular a section (700x700 pixels) of a scene acquired was selected. This image was assumed to be X1 image of the data set. The X2 image was artificially generated from the reference one. A first version of the X2 image was obtained by inserting some changes in the X1 image in order to simulate land cover variations. Then the histogram of the resulting image was slightly shifted to simulate different light conditions in the two images. Finally, two versions of the X2 image were generated by adding different realizations of zero-mean Gaussian noise to the X2 image (Signal-to-Noise-Ratio (SNR=20 and 10 dB). For simplicity, we assumed the spatial independence of the noise components in the images. As an example, Fig. 1(a) shows the band 4 of the reference image, X1. The map of the areas with simulated changes is presented in Fig. 1(b). For the two pairs of synthetic images considered, the corresponding difference images were obtained by applying the CVA technique. Spectral change detection techniques involve the transformation of two original images to a new single-band or multi-band images, in which the areas of spectral change are highlighted. For the two pairs of images considered, a single-band image was created based on the corresponding difference image. The difference images, used in pre-classification experiments, were obtained by applying the Change Vector Analysis (CVA) technique [5], [8], [10]. In order to produce spectral change data, only the magnitude value of difference multiband image vector was used. The threshold vectors were (62, 62) and (65,65) for the dataset with SNR=20 dB and 10 dB, respectively.



(a) (b)

Fig. 2. Synthetic data set utilized in the experiments. (a) Band 4 of the reference image, (b) map of the areas with simulated changes used as the reference map in the experiments.

In order to interpret classification accuracies we used two descriptive measures: the overall accuracy and the Kappa coefficient. The overall accuracy is computed by dividing the total number of correctly classified pixels by the total number of reference pixels. Kappa coefficient of agreement is a measure for overall thematic classification accuracy and ranges from 0 to 1. It is a measure of the difference between the actual agreement between reference data and an automated classifier and the chance agreement between the reference data and a random classifier. A true agreement (observed) approaches 1, and chance agreement approaches 0. A Kappa coefficient of 0 suggests that a given classification is no better than a random assignment of pixels [14].
Tables 1, and 2 show the confusion matrix for the datasets with SNR=20 dB and 10 dB, respectively.

Table 1. Confusion Matrix for the dataset with SNR = 20 dB.

	Not Changed (Reference)	Changed (Reference)
Not Changed (Classified)	479131	1939
Changed (Classified)	0	8930

Table 2. Confusion Matrix for the dataset with SNR = 10 dB.

	Not Changed (Reference)	Changed (Reference)
Not Changed (Classified)	479131	2064
Changed (Classified)	0	8805

Table 4 resumes the results and presents the overall accuracy and the Kappa coefficient for the two datasets. The worst performance was obtained by the dataset with SNR=10 dB, Kappa coefficient 0.89, and the best performance was obtained by the dataset with SNR=10 dB with a Kappa coefficient equal to 0.90.
We can note that resulted errors were only omission errors and the noise level did not affect significantly the performance of the algorithm.

Table 3. Results for two datasets.

	Kappa-coefficient	Overall Accuracy
SNR = 20 dB	0.90	99.60 %
SNR = 10 dB	0.89	99.57 %

5 Conclusions

The proposed method presents some improvements in the field of change detection and visualization of the certainty and magnitude of changes. In addition, a system to prioritize areas targeted for map and database revision based on a manager's criteria of a cost-effective threshold of change is presented. A 2-D fuzzy partition characterized by parameters a, and b is proposed which divides a 2-D histogram into two fuzzy subsets "changed" and "not changed." For each fuzzy subset, one fuzzy entropy and one non-fuzzy entropy are defined based on the fuzziness of the regions. The best fuzzy partition was found based on the maximum fuzzy entropy principle, and the corresponding parameters a, and b determines the fuzzy region [a; b]. The threshold is selected as the crossover point of the fuzzy region.
Further research should be conducted to test the potential improvements associated with such approach. Another selection of membership functions could be used, the possibility of using different parameters for the membership functions relative to X and Y sets could be experimented. In spite of the simplicity adopted, even the case characterized by high level of noise, the experimental results confirm the effectiveness of the presented technique.

References

1. Jensen, J. "Principles of change detection using digital. remote sensor data". In: Star, Estes, McGwire Eds. , Integration of Geographic Information Systems and Remote Sensing. Cambridge Univ. Press, pp. 37–54, 1997.
2. Johnson, R. D.; Kasischke, E. S. "Change Vector Analysis: a Technique for the Multispectral Monitoring of Land Cover and Condition", International Journal of Remote Sensing, No.19, pp. 411-426, 1998.

3. Malila, W. A., “Change Vector Analysis: An Approach for Detecting Forest Changes with Landsat”, Proceedings of the Machine Processing of Remotely Sensed Data Symposium, Ann Arbor, ERIM, pp. 326-335,1980.
4. Zhan, X.; Huang, C.; Townshend, R. de F.; Hansen, M.; Dimiceli, C.; Sholberg, R.; Hewson-Sacardeletti, J.; Tompkins, A. “Land Cover Detection with Change Vector in the Red and Near-Infared Reflectance Space”, Geoscience and Remote Sensing Symposium, 2000 IEEE International.
5. P. K. Sahoo, S. Soltani, A. K. C.Wong, and Y. Chen, “A survey of thresholding techniques,” Comput. Vis., Graph., Image Process., vol. 41, pp. 233–260, 1988.
6. H. D. Cheng, Y. H. Chen, and X. H. Jiang, “ Thresholding Using Two-Dimensional Histogram and Fuzzy Entropy Principle”, IEEE Transactions on Image Processing, Vol. 9, No. 4, pp. 732-735, 2000.
7. L. A. Zadeh, “Probability measures of fuzzy events,” J. Math. Anal. Applicat , vol. 23, pp. 421–427, 1968.
8. Leondes, C. T., “Image Processing and Pattern Recognition”, Academic Press, 1998.
9. Lunetta, R. S.; Elvidge, C. D., “Remote Sensing Change Detection”, Ann Arbor Press, 1998.
10. C. A. Murthy and S. K. Pal, “Histogram thresholding by minimizing graylevel fuzziness,” Inform. Sci., vol. 60, pp. 107–135, 1992.
11. Jensen, J., 1996. Introductory Digital Image Processing: A Remote Sensing Perspective, 2nd edn. Prentice-Hall, Englewood Cliffs, NJ, 330 pp.
12. Mathers, P. M., Computer Processing of Remotely-Sensed Images, John Wiley & Sons, 1999.
13. Chavez, P.S., Mackinnon, D.J. “Automatic Detection of Vegetation Changes in the Southwestern United States Using Remotely Sensed Images”, Photogrammetric Engineering and Remote Sensing, Vol 60, pp. 571-583, 1994.
14. Lillesand and Kieffer “Remote Sensing and Image Interpretation”, 4th Edition, John Wiley & Sons, Inc., 2000.