Unsupervised Change Detection Using Iterative Mixture Density Estimation and Thresholding

No-Wook Park, Kwang-Hoon Chi
Geoscience Information Center
Korea Institute of Geoscience and Mineral Resources
30 Gajeong-dong, Yuseong-gu, Daejeon 305-350, Korea
nwpark@kigam.re.kr

Abstract: We present two methods for the automatic selection of the threshold values in unsupervised change detection. Both methods consist of the same two procedures: 1) to determine the parameters of Gaussian mixtures from a difference image or ratio image, 2) to determine threshold values using the Bayesian rule for minimum error. In the first method, the Expectation-Maximization algorithm is applied for estimating the parameters of the Gaussian mixtures. The second method is based on the iterative thresholding that successively employs thresholding and estimation of the model parameters. The effectiveness and applicability of the methods proposed here are illustrated by an experiment on the multi-temporal KOMPAT-1 EOC images.
Keywords: Change Detection, Gaussian Mixtures, Threshold.

1. Introduction

Change detection using multi-temporal data is one of the most important applications using remote sensing data. The development of effective analysis algorithms for change detection provides the possibility of solving complex problems related to the Earth's environment monitoring [1].

Until now, various schemes for change detection have been proposed and applied to various application fields [2]. Especially, the unsupervised approach can simply identify the amount of change by making a direct comparison of two multi-temporal images. Though this approach provides no information on the nature of the change (i.e. 'from-to' change information), it has been widely used in many applications in which the ground truth is not available. Commonly used unsupervised change detection techniques include image differencing, image ratioing, change vector analysis, principal component analysis, etc.

In the unsupervised approach, some important issues arise such as radiometric calibration, precise geometric rectification, and the selection of the threshold values. Among these issues, we will focus on the selection of threshold values. Since the unsupervised approach only provides information on change or non-change, the accuracy of the change detection result depends on the selection of the threshold value between changed and unchanged pixels. Traditionally, the standard deviation multiplied by a factor in the difference or ratio image has been commonly used to define thresholds [3]. However, the selection of the multiplying factor is based on empirical trial-and-error procedure. So when we have no information on the ground truth, the accuracy of above approach heavily depends on the analyst's subjective criteria, which may lead to unreliable change detection results [4].

To overcome these drawbacks, we present the methods for the automatic selection of the threshold values. We assume that the difference or ratio image can be modeled as Gaussian mixtures. Once the Gaussian mixture density model has been determined, the optimal threshold value is determined by using the Bayesian rule for minimum error. Under the assumption of the Gaussian mixtures, two different iterative methods will be applied (Fig. 1).

In the following sections, the details of the theory and procedures of the proposed methods are described. Application results of the method to remote sensing images are also shown.

2. Theory and Procedures

1) Gaussian Mixtures

In this paper, we assume the density distribution of the difference or ratio image can be represented as a linear combination of Gaussian component densities associated with the changed and unchanged pixels (Fig. 2). In [4], they proposed a two-component Gaussian mixture density model in the change vector image. In the difference or ratio image, however, the changed areas generally consist of two areas: areas where the DN values are decreased or increased. So we extend their model to the three-component Gaussian mixture density model (i.e. positive change, negative change, non-change).

Let X be a random variable in the difference or ratio image. If we assume the Gaussian mixtures, the probability density function $p(X)$ can be represented as:

$$p(X) = p(\mathbf{w}_{c+})\,p(X \mid \mathbf{w}_{c+}) + p(\mathbf{w}_{c-})\,p(X \mid \mathbf{w}_{c-}) + p(\mathbf{w}_{nc})\,p(X \mid \mathbf{w}_{nc}) \quad (1)$$

where $\mathbf{w}_{c+}$, $\mathbf{w}_{c-}$ and $\mathbf{w}_{nc}$ denote positively changed areas, negatively changed areas and unchanged areas, respectively.

For the Gaussian mixtures, each component density is a normal probability distribution:

$$p(X \mid \mathbf{w}_i) = \frac{1}{\mathbf{s}_i \sqrt{2\mathbf{p}}} \exp[-\frac{(X - \mathbf{m}_i)^2}{2\mathbf{s}_i^2}] \quad (2)$$

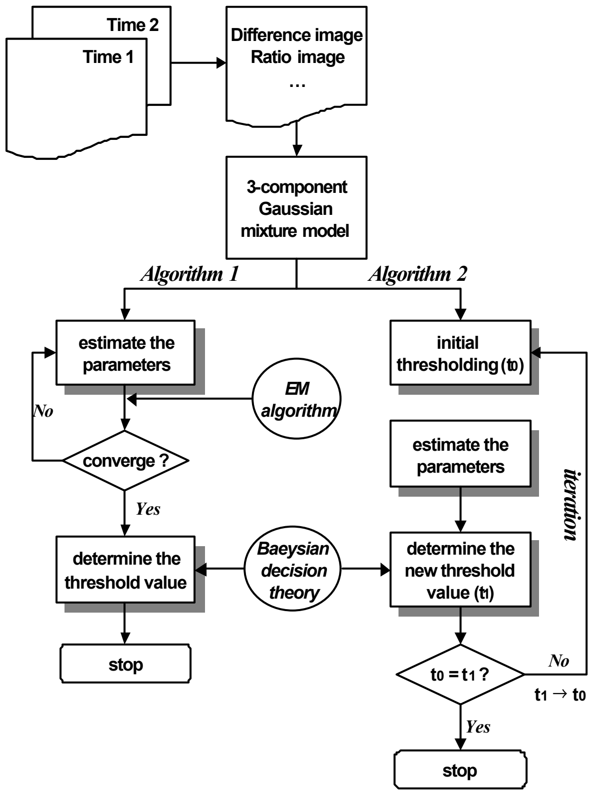

Fig. 1. Two algorithms for the selection of threshold values.

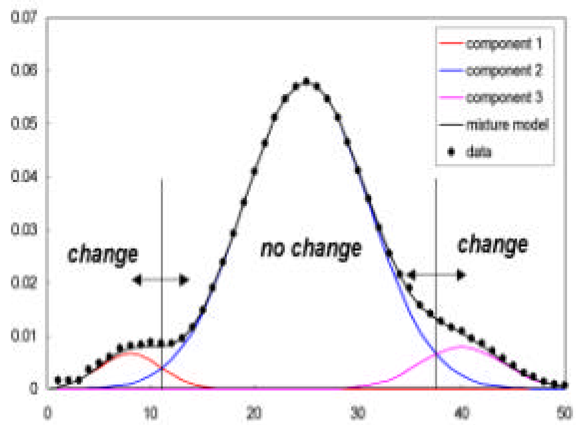

Fig. 2. A three-component Gaussian mixture model.

where m_i and s_i denote the mean and standard deviation of w_i, respectively.

2) Bayesian Rule for Minimum Error

If the model parameters of the Gaussian mixtures were estimated, we select optimal threshold values by using the Bayesian rule for minimum error [5].

According to the Bayesian rule for minimum error, optimal thresholds X are determined as the appropriate solution of

$$\begin{aligned} p(w_{nc})p(X \mid w_{nc}) &= p(w_{c+})p(X \mid w_{c+}) \\ p(w_{nc})p(X \mid w_{nc}) &= p(w_{c-})p(X \mid w_{c-}) \end{aligned} \tag{3}$$

Those pixels that have DN value X satisfying inequality $p(w_{nc})p(X \mid w_{nc}) > p(w_{c+})p(X \mid w_{c+})$ will be classified into w_{nc}, otherwise into w_{c+}. Eq. (3) guarantees the smallest misclassification error.

3) Algorithm I

The algorithm I implements iterative estimation of the model parameters and then determines the threshold values.

Various methods have been developed for determining the model parameters of Gaussian mixtures from the data set. In this paper, we adopt the EM algorithm [6] that iteratively modifies the parameters of Gaussian mixtures to maximize the likelihood of the data.

The EM algorithm consists of two major steps: the expectation step, followed by the maximization step. In the expectation step, we do a soft assignment of each observation to each Gaussian component model. The maximization step then provides a new estimate of the parameters. These two steps are iterated until convergence.

After determining the model parameters using the EM algorithm, the threshold values are finally determined using Eq. (3).

4) Algorithm II

Unlike the algorithm I, the algorithm II iteratively estimates the model parameters and the threshold values.

First, the difference or ratio image is initially thresholded at some DN values and then the model parameters of the thresholded image are determined. Having applied Eq. (3), a new threshold value is estimated. The new threshold value is then applied for thresholding the image. This procedure is iteratively implemented until no further change in the threshold value occurs.

3. Experiments

To assess the effectiveness of the proposed methods, we applied the methods to a multi-temporal KOMPSAT-1 EOC data set. The data set consists of two panchromatic images acquired by the KOMPAT-1 EOC sensor in the western part of Daejeon, Korea in March 2001 and May 2001, respectively. The available ground truth concerning the changed areas was used to assess the change detection errors. This reference map was refined by a manual analysis of the images considered. The area selected for the experiment consisted of 500 by 500 pixels and the number of the changed areas was 10,513.

Before applying the change detection method, image normalization based on the multiple regression technique was implemented. Then the image differencing technique was applied to the data set.

Table 1. Comparison of true values and estimates and threshold values obtained by the two algorithms .

		Algorithm I	Algorithm II	True values
Mean	(-) change	-30.669	-30.479	-30.728
	Non-change	-0.502	-0.502	-0.503
	(+) change	19.075	19.284	18.900
Standard deviation	(-) change	22.311	22.231	22.701
	Non-change	1.099	1.105	1.101
	(+) change	11.902	11.969	11.823
Prior probability	(-) change	0.003	0.003	0.003
	Non-change	0.959	0.962	0.958
	(+) change	0.038	0.035	0.039
Threshold	(-) change ~ non-change	-4.28	-4.30	
	Non-change ~ (+) change	4.42	4.46	

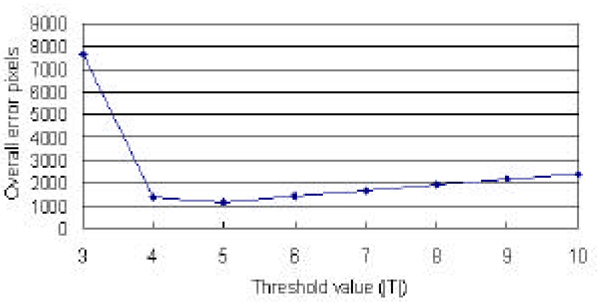

Fig. 3. Variation of overall errors with respect to the threshold values.

The two algorithms for automatic selection of threshold values were applied separately to the difference image. Firstly, we compared the parameter values estimated by the two algorithms with those of the ground truth data (Table 1). The estimates from the algorithm I were converged after 6 iterations and for the algorithm II after 5 iterations.

As shown in Table 1, the two proposed algorithms provided estimates of the parameters very close to the corresponding true values, though a priori probability of the unchanged areas estimated by the algorithm II was slightly different from the real values. Finally, the two algorithms showed the same threshold values. As a result, those pixels where the absolute values of the DN values were greater than four (i.e. $|X|$ 5) were classified into the changed pixel. Since the accuracy of the proposed algorithms depends on the accuracy of the estimates of the parameters in the Gaussian mixtures, this result would have incurred from the similarities of the estimated model parameters.

To evaluate the change-detection accuracy obtained from the two proposed methods, we compared the threshold values obtained by the proposed algorithms with the minimum error threshold value derived by the empirical trial-and-error procedure. From Fig. 3, we can see that the threshold value was the same as the minimum error threshold value (i.e. ±5).

4. Conclusions

In this paper, two methods for the automatic selection of the threshold values in unsupervised change detection have been presented and applied. The applicability and validity of the proposed methods were demonstrated by the experiment using multi-temporal remote sensing images. Unlike traditional methods, the two methods proposed here could determine the threshold values in an unsupervised manner under the assumption of the Gaussian mixtures. The algorithm II showed the result similar to the extended version of the algorithm I proposed by [4]. The determined threshold values also corresponded to the values showing the minimum overall change-detection errors. These experiment results confirm that the methods can be effectively applied to unsupervised change detection in case that we have no ground truth information in the study area of interest.

Acknowledgement

This research was supported by the Korean Ministry of Science and Technology. The authors thank Mr. K.J. Lee of KARI, Korea for providing the KOMPSAT-1 EOC data set.

References

[1] Bruzzone, L. and P. Smits, 2001. *Analysis of multi-temporal remote sensing images*, World Scientific, NJ.

[2] Singh, A., 1989. Digital change detection techniques using remotely-sensed data, *Int. J. Remote Sensing*, 10(6): 989-1003.

[3] Jensen, J.R., 1996. *Introductory digital image processing*, Prentice Hall, NJ.

[4] Bruzzone, L. and D.F. Prieto, 2000. Automatic analysis of the difference image for unsupervised change detection, *IEEE Trans. Geosci. Remote Sensing*, 38(3):1171-1182.

[5] Fukunaga, K., 1990. *Introduction to statistical pattern recognition*, Academic Press, London, U.K.

[6] Dempster, A.P., N.M. Laird, and D.B. Rubin, 1977. Maximum likelihood from incomplete data via the EM algorithm, *J. Roy. Statist. Soc. Ser. B*, 39(1): 1-38.