

A Novel approach for unsupervised change detection in multi spectral images

Ms.V.Vidhya

Assistant Professor

Department of Computer Science & Engineering
Annai Vailankanni College of Engineering
vidhya.v@gmail.com

Mr.G.Jeyaram

Assistant Professor

Department of Computer Science & Engineering
Annai Vailankanni College of Engineering
jeyaramgj@gmail.com

Mrs.K.R.Ishwarya

Assistant Professor

Department of Computer Science & Engineering
Annai Vailankanni College of Engineering
yazhini1722011@gmail.com

Mrs.R.Hema

Assistant Professor

Department of Computer Science & Engineering
Annai Vailankanni College of Engineering
hema25me@gmail.com

ABSTRACT

The paper a novel approach for unsupervised change detection in multi spectral images presents a novel approach to unsupervised change detection in multispectral remote-sensing images. The proposed approach aims at extracting the change information by jointly analyzing the spectral channels of multitemporal images in the original feature space without any training data. This is accomplished by using a selective Bayesian thresholding for deriving a pseudo training set that is necessary for initializing an adequately defined binary semisupervised support vector machine (S^3VM) classifier. Starting from these initial seeds, the S^3VM performs change detection in the original multitemporal feature space by gradually considering unlabeled patterns in the definition of the decision boundary between changed and unchanged pixels according to a semisupervised learning algorithm. The values of the classifier parameters are then defined according to a novel unsupervised model-selection technique based on a similarity measure between change-detection maps obtained with different settings. Experimental results obtained on different multispectral remote-sensing images confirm the effectiveness of the proposed approach.

1. INTRODUCTION

The ever increasing number of operational satellites for Earth observation results in a growing interest of the remote-sensing community in the analysis of images acquired on the same geographical area at different times.

Depending on the specific problem addressed, the analysis of multitemporal data can be carried out according to supervised classification techniques or to unsupervised change-detection procedures. The methods adopted for data analysis also depend on the kind of images considered (e.g., multispectral or SAR) and on the related properties. This paper focuses the attention on unsupervised change detection in multispectral images.

The main novel contributions of this paper are as follows: i) the basic idea (which consists in the synergistic exploitation of both the simplicity associated with the thresholding of the magnitude of SCVs and the completeness of information contained in the original feature space) and the definition of the architecture of the change-detection approach; and ii) the procedure proposed for unsupervised S^3VM classification [1].

1.1. CHANGE DETECTION

Change detection is the process of automatically identifying analyzing and detecting regions that have undergone changes from a multi-spectral and multi-temporal images. The detection of changes that occur on the Earth surface by using multitemporal remote sensing images is one of the most important applications of the remote sensing technology. Remote sensing data are primary sources extensively used for change detection in recent decades. Many change detection techniques have been developed [9]. The image change results from a combination of underlying factors

- Appearance or disappearance of objects
- Motion of objects relative to background
- Shape changes of objects
- Changes in brightness or color

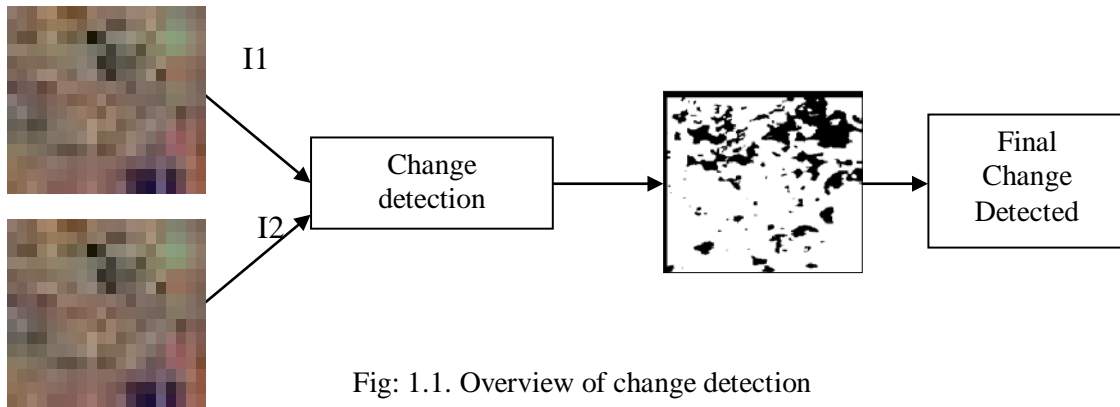


Fig: 1.1. Overview of change detection

Good change detection research should provide the following information: (1) area change and change rate; (2) spatial distribution of changed types; (3) change trajectories of land-cover types; and (4) accuracy assessment of change detection results. When implementing a change detection project, three major steps are involved: (1) image preprocessing including geometrical rectification and image registration, radiometric and atmospheric correction, and topographic correction if the study area is in mountainous regions; (2) selection of suitable techniques to implement change detection analyses; and (3) accuracy assessment.

1.2. SELECTION OF THRESHOLDS

Two methods are often used for selection of thresholds:

(1) Interactive procedure or manual trial-and-error procedure—an analyst interactively adjusts the thresholds and evaluates the resulting image until satisfied; and

(2) Statistical measures— selection of a suitable standard deviation from a class mean.

2. IMAGE CHANGE DETECTION

Let us consider I_1 and I_2 be two co registered multispectral input images of size $P \times Q$ acquired on the same geographical area at different times t_1 and t_2 , respectively. Let D be the number of spectral channels of each image, and let $\Omega = \{\omega_u, \omega_c\}$ be the set of classes of unchanged and changed pixels to be

identified. Then the proposed technique consists of three main parts:

1. An initialization that exploits a selective Bayesian thresholding of the magnitude of SCVs for generating a set of seed pixels (pseudotraining set).
2. A properly designed S3VM classifier that exploits the pseudotraining set and the unlabeled pixels of images I_1 and I_2 for producing the change-detection map by analyzing the original multidimensional feature space.
3. A novel similarity measure that is applied to the change-detection maps obtained with S³VM classifier in an unsupervised way.

2.1. BAYESIAN THRESHOLDING

Unlike classical unsupervised methods used in remote sensing applications, our approach involves formulating the problem of the analysis of the difference image for change detection in terms of the Bayesian decision theory. Let $I = \{X_p\}_{p=1}^{P \times Q}$ be the set of $P \times Q$ d -dimensional vectors (where $d = 2 \times D$) whose components are the channels associated with the generic p th pixel in I_1 and I_2 . I represent the multitemporal data set under analysis achieved by stacking the co registered multispectral images I_1 and I_2 .

BLOCK DIAGRAM

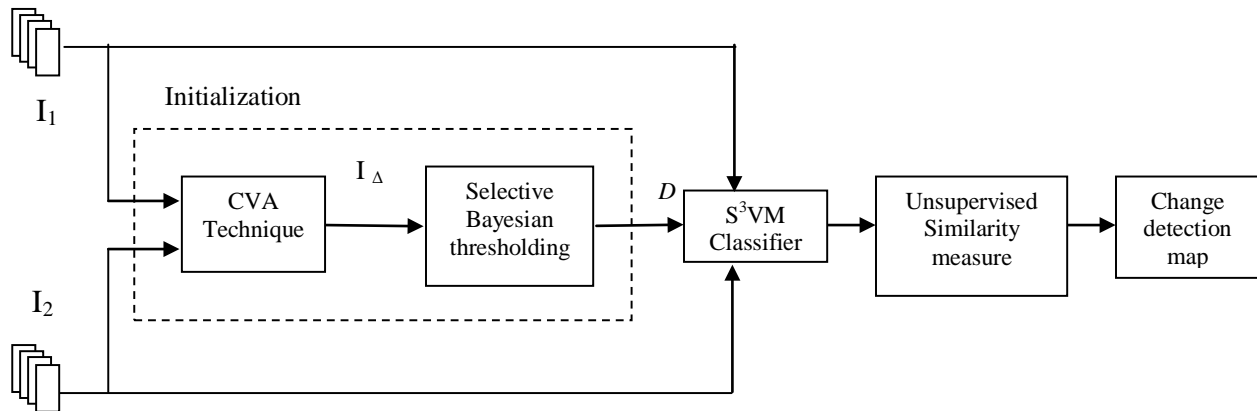


Fig: 2.1 Overview of the project

The first step of the proposed unsupervised approach to change detection aims at identifying a pseudotraining set $D = \{X_l, Y_l\}$ made up of pairs (x_n^l, y_n^l) to be used as seed patterns for initializing the S^3VM , where $X_l = \{x_n^l, x_n^l \in I\}$ and $Y_l = \{y_n^l, y_n^l \in \Omega\}$ (Y_l is the set of pseudo labels corresponding to seed patterns in X_l). This subset should contain patterns that are associated either to changed or unchanged areas with no uncertainty. As we detect these patterns in an automatic and unsupervised way, it is assumed that patterns included in the subset X_l are associated with a high probability to belong to changed or unchanged areas. Accordingly, we propose to identify X_l by applying the CVA technique to I_1 and I_2 and by selectively thresholding the magnitude I_Δ of SCVs [6].

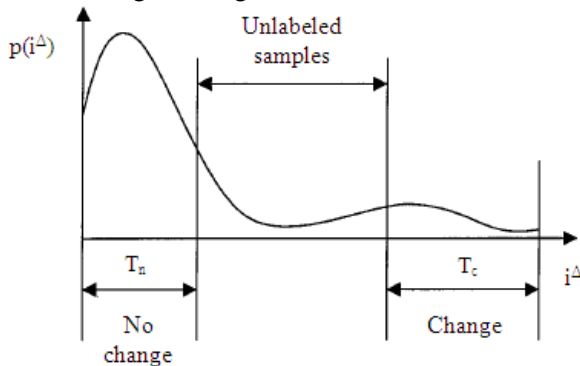


Fig: 2.2. Schematic representation of the thresholding strategy applied to the difference-image for the initialization of the EM algorithm.

2.2. EXPECTATION-MAXIMIZATION (EM) ALGORITHM

In order to analyze the difference image on the basis of the Bayes theory, the main problems to be solved are the estimations of both the probability density functions and the *a priori* probabilities of the classes. However, as we deal with an unsupervised approach, the estimation process cannot be performed on the basis of a training set. This project proposes an unsupervised method for estimating the aforesaid statistical terms. In particular, the method assumes that the probability density function $p(X)$ computed on the pixel values in the difference image I_Δ can be modeled as a mixture density distribution consisting of two density components associated with the classes ω_n and ω_c , respectively, i.e.,

$$p(X) = p(X/\omega_n)P(\omega_n) + p(X/\omega_c)P(\omega_c) \quad (2.2)$$

Under this assumption, the unsupervised estimates of $p(X/\omega_n)$, $p(X/\omega_c)$, $P(\omega_n)$ and $P(\omega_c)$ can be performed by using the EM algorithm [6].

The EM algorithm consists of an expectation step and a maximization step, which are iterated until convergence. The expectation step is computed with respect to the unknown underlying variables, using the current estimates of the parameters, and is conditioned by observations. The maximization step provides new estimates of the parameters. We propose using the EM algorithm to estimate the values of the *a priori* probabilities and the values of the parameters that characterize the density functions. Let us assume that both can be modeled by Gaussian distribution. The Gaussian distribution is given by equation (2.3).

$$G = \frac{1}{\sigma\sqrt{2\Pi}} e^{-(x-\mu)^2/2\sigma^2} \quad (2.3)$$

Initially before iteration starts the prior probabilities are calculated using the following formula (2.4) and (2.5)

$$P(\omega_c) = \frac{n(\omega_c)}{n(\omega_c) + n(\omega_n)} \quad (2.4)$$

$$P(\omega_n) = \frac{n(\omega_n)}{n(\omega_c) + n(\omega_n)} \quad (2.5)$$

Where $n(\omega_c)$ =Total number of changed pixels and $n(\omega_n)$ =total number of unchanged pixels.

When the iteration starts the density function associated with the class ω_n can be described by the mean μ_n and the variance σ_n . Analogously, the density function associated with the class ω_c can be described by the mean μ_c and the variance σ_c . It is possible to prove that the equations for estimating the aforementioned statistical terms for the class of not changed pixels $P(\omega_n)$ are the following

$$P^{t+1}(\omega_n) = \sum_{X(i,j) \in I_\Delta} \frac{P^t(\omega_n) p^t(X(i,j) | \omega_n)}{p^t(X(i,j))} \quad (2.6)$$

$$\mu_n^{t+1} = \frac{\sum_{X(i,j) \in I_\Delta} \frac{P^t(\omega_n) p^t(X(i,j) | \omega_n)}{p^t(X(i,j))} X(i,j)}{\sum_{X(i,j) \in I_\Delta} \frac{P^t(\omega_n) p^t(X(i,j) | \omega_n)}{p^t(X(i,j))}} \quad (2.7)$$

$$(\sigma_n^2)^{t+1} = \frac{\sum_{X(i,j) \in I_\Delta} \frac{P^t(\omega_n) p^t(X(i,j) | \omega_n)}{p^t(X(i,j))} [X(i,j) - \mu_n^t]^2}{\sum_{X(i,j) \in I_\Delta} \frac{P^t(\omega_n) p^t(X(i,j) | \omega_n)}{p^t(X(i,j))}} \quad (2.8)$$

where the superscripts t and $t+1$ denote the values of the parameters at the current and next iterations, respectively. Analogous equations are used to estimate the prior probability and the mean and variance values of the conditional density function associated with the class ω_c .

In particular, a subset S_n of pixels likely to belong to ω_n and a subset S_c of pixels likely to belong to ω_c can be obtained by applying two thresholds, T_n and T_c , calculated using equations (2.9) and (2.10), to the right and left extremes of the histogram of the difference image.

$$T_n = M_D(1 - \alpha) \quad (2.9)$$

$$T_c = M_D(1 + \alpha) \quad (2.10)$$

where M_D is the middle value of the histogram, and $\alpha \in (0,1)$ is the initialization parameter that defines the range around M_D in which pixels cannot be easily identified as either changed or unchanged. M_D can be calculated using the formula

$$M_D = [\max \{I_\Delta\} - \min \{I_\Delta\}] \quad (2.11)$$

Then the sets given by equations (2.12) and (2.13)

$$S_n = \{X(i,j) | X(i,j) < T_n\} \quad (2.12)$$

$$S_c = \{X(i,j) | X(i,j) > T_c\} \quad (2.13)$$

are used to compute the initial estimates of the statistical parameters associated with the classes ω_n and ω_c , respectively.

The previously described formulation of the EM algorithm allows one to estimate the statistical parameters associated with both classes of ω_n and ω_c under the assumption of Gaussian distributions. The estimates obtained by the EM algorithm at convergence can be exploited to analyze the difference image in the coming steps.

In the Gaussian case, the optimal threshold is equivalent to solving the following quadratic equation (2.14):

$$(\sigma_n^2 - \sigma_c^2)T_0^2 + 2(\mu_n \sigma_c^2 - \mu_c \sigma_n^2)T_0 + \mu_c^2 \sigma_n^2 - \mu_n^2 \sigma_c^2 - 2\sigma_n^2 \sigma_c^2 \ln \left[\frac{\sigma_c P(\omega_n)}{\sigma_n P(\omega_c)} \right] = 0 \quad (2.14)$$

The precision of the final change-detection map depends on the accuracies of the estimates provided by the EM algorithm.

2.3. SEMI SUPERVISED SUPPORT VECTOR MACHINE (S³VM)

The main idea of the second step of the proposed technique is to define a discriminate function in the original multitemporal feature space of the multitemporal images, which can accurately separate changed pixels from unchanged ones. In this way, it is possible to integrate the incomplete seed information extracted in the previous phase with the complete information associated with the uncertain unlabeled pixels in the original multitemporal images.

In the presence of a reliable training set, the aforementioned problem could be easily solved according to any binary classification technique applied to the multitemporal feature space. However, in our case, we have a pseudotraining set which is

made up of samples that satisfy a tradeoff between high probability to be correctly labeled and complete representation of the problem. In other terms, the margin imposed in the thresholding procedure used in the first step results in a pseudotraining set that does not model completely the statistics of changed and unchanged classes. In particular, we expect that the pseudotraining set does not contain enough information for modeling the distribution of patterns that are close to the decision boundary [1].

The proposed S3VM technique is based on the following two main phases:

- 1) Initialization and
- 2) Semisupervised Learning.

2.3.1. PHASE 1—INITIALIZATION

Let $D^{(i)} = \{X^{(i)}, Y^{(i)}\}$ and $X^{*(i)}$ denote the pseudotraining set and the unlabeled set (i.e., the set containing all the samples that are not associated with a label yet) at the generic iteration i , respectively. The first phase corresponds to the initial step of the entire process ($i = 0$). We have that $D^{(0)} = \{X^{(0)}, Y^{(0)}\} \equiv \{X, Y\}$ and $X^{*(0)} \equiv X_u$. The learning cost function of a standard semi supervised SVM is used to obtain an initial separation hyperplane based only on pseudotraining data [4].

The standard semi supervised SVM is a linear inductive learning classifier, where the data in the input space are separated by the hyperplane as in equation (2.15)

$$y(X) = w^T \cdot X + b \quad (2.15)$$

with the maximal geometric margin

$$\varphi(w) = \frac{2}{w^T \cdot w} \quad (2.16)$$

where w is a vector normal to the hyperplane, b is a constant such that $b/\|w\|$ represents the distance of the hyperplane from the origin, and T denotes the transpose of a vector. The minimal hyperplane is given by (2.17)

$$\varphi(w) = \min_w \left\{ \frac{1}{2} (w^T \cdot w) \right\} \quad (2.17)$$

To allow some training errors for generalization, the slack variables ξ_l and the associated regularization parameter C are introduced for the nonseparable cases

$$\varphi(w) = \min_{w, \xi_l} \left\{ \frac{1}{2} (w^T \cdot w) + C \sum_{l=1}^n \xi_l \right\} \quad (2.18)$$

$$\text{s.t.} \left\{ \begin{array}{l} \forall_{l=1}^n : Y_l (w^T \cdot X_l + b) \geq 1 - \xi_l \\ \forall_{l=1}^n : \xi_l > 0. \end{array} \right\}$$

The training samples associated with the nonzero multipliers (called support vectors) contribute to defining the separating hyperplane. If the data in the input space cannot be linearly separated, they are projected into a higher dimensional feature space with a nonlinear mapping function $\varphi(x)$. According to standard SVM notation, the labels ω_u and ω_c are represented with values “+1” and “-1,” respectively (i.e., $y_n^l \in \{\omega_c, \omega_u\} \equiv \{+1, -1\}$).

2.3.2. PHASE 2—SEMISUPERVISED LEARNING

The second phase of the proposed S³VM starts with iteration $i = 1$ and represents the core of the algorithm. At the generic iteration i , pseudolabels are given to unlabeled pixels belonging to $X^{*(i)}$ X_u according to the current separation hyperplane. These pixels are called pseudolabeled patterns. As support vectors are the only patterns that affect the position of the discriminant hyperplane, unlabeled samples which fall into the margin and are closest to the margin bounds have the highest probability to be correctly classified.

Accordingly, the ρ pseudolabeled samples lying into the margin that are closest either to the lower or the upper margin bound are selected and denoted as semilabeled patterns. The set containing all the semilabeled samples defined at iteration i is called $H^{(i)}$. Patterns of $H^{(i)}$ and their corresponding semilabels are then merged with $X^{(i)}$ and $Y^{(i)}$, respectively. Let $J^{(i)}$ is calculated as in equation (2.19) which represent the set of all the pixels selected from X_u , which have been always assigned the same label until iteration i .

$$J^{(i)} = (J^{(i-1)} - S^{(i)}) \cup H^{(i)} \quad (2.19)$$

$$X^{*(i)} = (X^{*(i-1)} - H^{(i-1)}) \cup S^{(i-1)} \quad (2.20)$$

A dynamical adjustment is necessary for considering the modifications in the position of the separation hyperplane. Let $S^{(i)}$, represent the set of samples belonging to $J^{(i-1)}$ whose labels obtained according to the separation hyperplane at iteration i are different than those at iteration $i - 1$ (*label inconsistency*). Patterns belonging to $S^{(i)}$ (2.21) are reset to the unlabeled state and moved again into $X^{*(i)}$ (2.20). In this way, it is possible to reconsider these patterns at the following iterations.

$$S^{(i)} = \begin{cases} \phi, i = 0,1 \\ \{X^u \mid (X^u \in X^{(i)}, x^u \in X^{(i-1)}), y^{u(i)} \neq \dots \end{cases} \quad (2.21)$$

The semisupervised learning is iterated until convergence, which is obtained when none of the originally unlabeled samples lies into the margin band.

2.4. SIMILARITY MEASURE

In the proposed strategy, the most reliable solution (change detection map) is selected according to the following criteria.

- 1) Among all the considered solutions, select those that result in an accuracy on the pseudotraining set *D*.
- 2) The accuracy is calculated using percentage of correct classification

The solution that shows the highest average similarity with all the others is selected as the output of the proposed change-detection technique. This solution is the most regular among the considered ones and, thus has the highest probability to be associated with a model that properly seizes the properties of the change-detection problem.

It is worth noting that the aforementioned criteria result in a reliable model selection if the pseudotraining set contains samples that represent all kinds of changes present in the images. As the EM-based thresholding algorithm is used for defining the pseudotraining set proved to be reliable and accurate in many change detection problems, also in the presence of different kinds of changes.

3. SIMULATIONS

INPUT IMAGES

Input image 1 and 2 are two multispectral and mutitemporal input image taken at same scene. These two images are given as input for detecting changes.

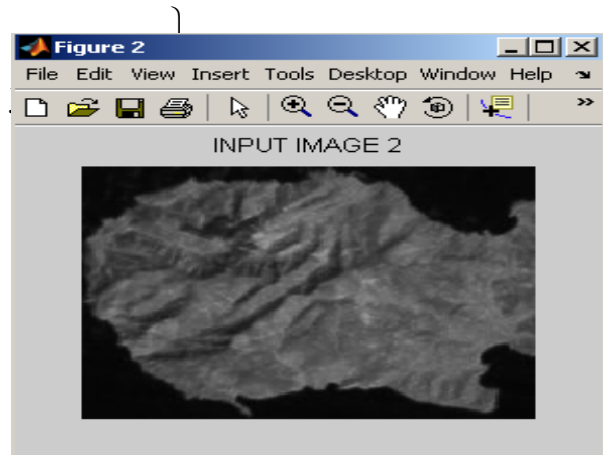
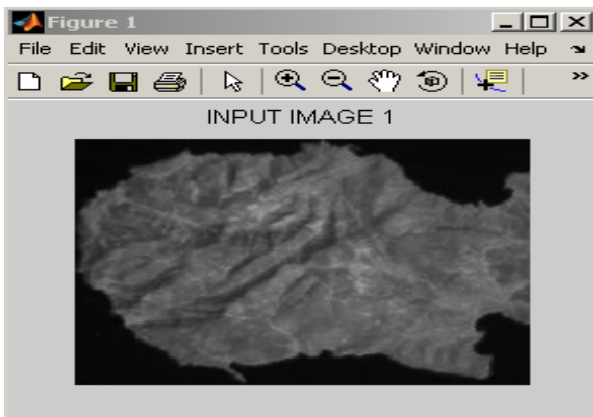


Fig: 3.2. Input images

DIFFERENCE IMAGE

The difference image obtained is the intermediate output obtained by finding the difference of the two input images 1 and 2. This difference image is further used for calculating the threshold value using Bayesian thresholding.

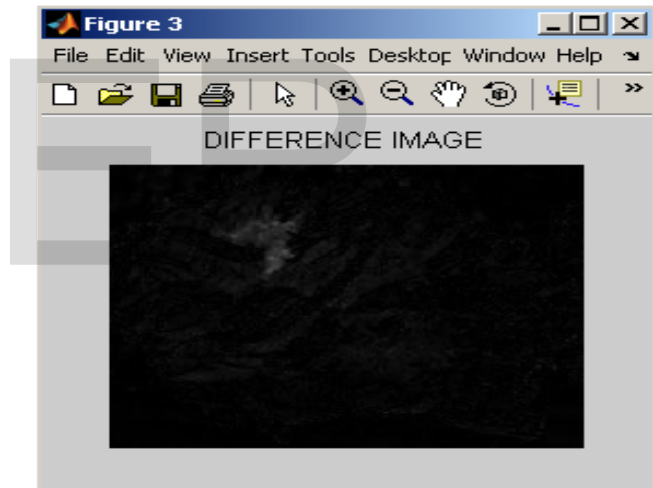


Fig: 3.3. Difference image

THRESHOLD CALCULATION

This graph plotted between the gray scale value and the probability density function (pdf) helps in finding the threshold value. The blue curve represents the histogram of the difference image obtained in the previous step. The red curve represents the unchanged pixels and the green curve represents the changed pixels. The intersection points of the red and green curves indicate the threshold value.

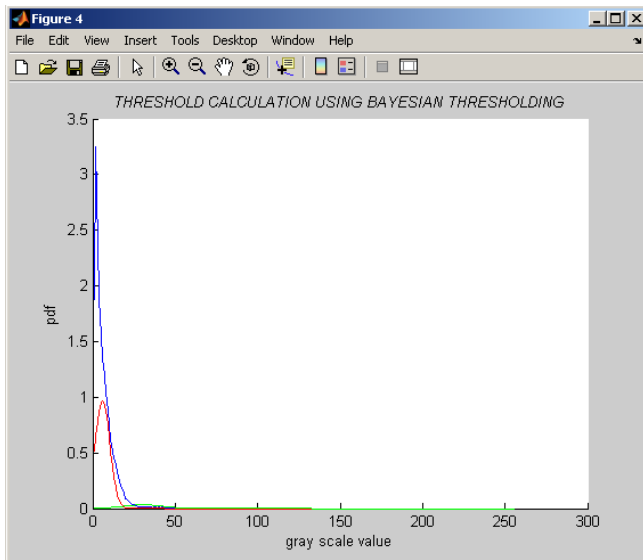


Fig: 3.4. Threshold calculation

DIFFERENCE IMAGE AFTER APPLYING THRESHOLD

The threshold calculated is applied and the difference is calculated again in order to extract the seed pixel for the S^3VM classifier.

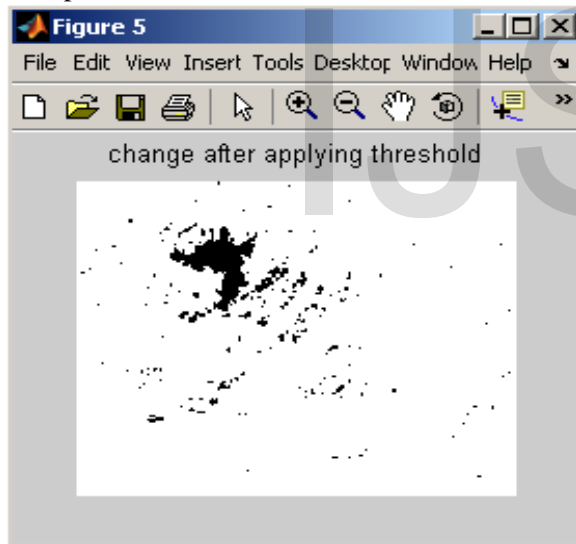


Fig: 3.5. Change after applying threshold

SEMISUPERVISED SUPPORT VECTOR MACHINE CLASSIFIER

The S^3VM classifier is trained using the extracted seed pixel, and then the unlabeled and semilabeled pixels are classified using S^3VM classifier. The green color indicates the changed pixels, red indicates the unchanged pixels and the blue color indicates the semilabeled pixels.

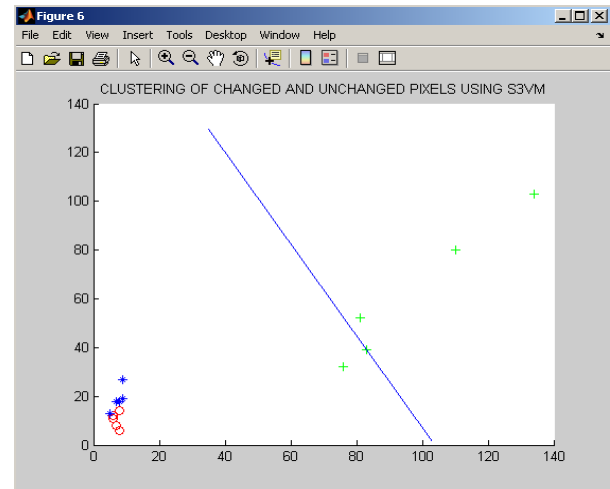


Fig: 3.6. S^3VM classifier

DIFFERENCE IMAGE AFTER APPLYING S^3VM CLASSIFICATION

This image in fig 3.7 is the final difference image obtained after S^3VM classification.

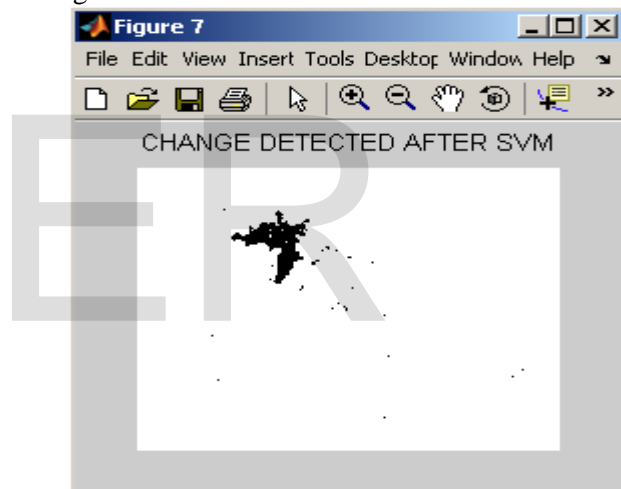


Fig: 3.7. Change detected after S^3VM classification

4. CONCLUSION

The proposed architecture properly exploits the information present in the original images by jointly analyzing the bitemporal spectral features. The initialization seed extracted by the Bayesian thresholding of the SCV magnitude and the joint use of unlabeled patterns has increased the accuracy of the decision boundary between changed and unchanged classes in a completely unsupervised way. Experimental results, obtained on two different remote sensing data sets, confirm the effectiveness of the proposed technique, the direct use of the spectral features of the considered images, sharply increased the change-detection accuracy with respect to the CVA algorithm applied with the EM-based thresholding. The change detected using the proposed technique has high accuracy.

REFERENCES

1. Francesca Bovolo, Lorenzo Bruzzone, and Mattia Marconcini (2008) “A Novel Approach to Unsupervised Change Detection Based on a Semisupervised SVM and a Similarity Measure,” *IEEE transactions on geoscience and remote sensing*, vol. 46, no. 7.
2. S. Ghosh, L. Bruzzone, S. Patra, F. Bovolo, and A. Ghosh, (2007) “A context-sensitive technique for unsupervised change detection based on Hopfield-type neural networks,” *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, pp. 778–789.
3. N. Cristianini and J. Shawe-Taylor (2000), “*An Introduction to Support Vector Machines*”. Cambridge, U.K.: Cambridge Univ. Press.
4. L. Bruzzone, M. Chi, and M. Marconcini, (2006) “A novel transductive SVM for the semisupervised classification of remote-sensing images,” *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 11, pp. 3363–3373.
5. F. Bovolo and L. Bruzzone, (2007) “A split-based approach to unsupervised change detection in large-size multitemporal images: Application to tsunami damage assessment,” *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 6, pp. 1658–1670.
6. L. Bruzzone and D. Fernández Prieto, (2000) “Automatic analysis of the difference image for unsupervised change detection,” *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 3, pp. 1171–1182.
7. L. Bruzzone and S. B. Serpico, (1997) “Detection of changes in remotely sensed images by the selective use of multi-spectral information,” *Int. J. Remote Sens.*, vol. 18, no. 18, pp. 3883–3888.
8. D. Lu, P. Mausel, E. Brondízio, and E. Moran, (2004) “Change detection techniques,” *Int. J. Remote Sens.*, vol. 25, no. 12, pp. 2365–2407.
9. R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, (2005) “Image change detection algorithms: A systematic survey,” *IEEE Trans. Image Process.*, vol. 14, no. 3, pp. 294–307.
10. P. L. Rosin, (2002) “Thresholding for change detection,” *Comput. Vis. Image Underst.*, vol. 86, no. 2, pp. 79–95.
11. V.N. Vapnik, (1998) “*Statistical Learning Theory*”. Hoboken, NJ: Wiley.