NOVEL METHOD FOR UNSUPERVISED FUZZY CHANGE DETECTION IN MULTISPECTRAL REMOTELY SENSED IMAGES

Arash Abadpour

Sharif University of Technology
Mathematics Science Department
P.O. Box 11365-9517, Tehran, Iran
abadpour@math.sharif.edu

Shohreh Kasaei

Sharif University of Technology
Computer Engineering Department
P.O. Box 11365-9517, Tehran, Iran
skasaei@sharif.edu

ABSTRACT

Analysis of the multi–spectral remotely sensed images of the areas destructed by an earthquake is proved to be a helpful tool for construction assessments. In this paper, we propose a new fast and reliable fuzzy change detection method for multi–spectral images. The proposed fuzzy change detection method is mathematically and experimentally investigated and shown to be efficient and effective.

Keywords: Remote Sensing, Multispectral Image Analysis, Change Detection, Fuzzy Clustering.

1. INTRODUCTION

In recent years, the spatial and the spectral resolution of the remote–sensing sensors and the revisiting frequency of the satellites, has been extensively increased. These developments, has offered the possibility of addressing new applications of remote–sensing in environmental monitoring [2, 6]. On the other hand, the officials are getting more and more aware of using multi–temporal remotely sensed images for regular and efficient control of the environment [25, 9]. Also, there are reports about the application of remote sensing change detection for global weaponry control (*e.g.*, see [7]).

A key issue in remote–sensing image analysis is to detect changes of the earth surface in order to manage possible interventions to avoid massive environmental problems [4]. Recently, many researchers have worked on using the remote–sensing data to help estimating the earthquake damage [15, 3] or the afterwards reconstruction progress [23, 17].

Change detection algorithms usually take two images as the two shots before and after the change and return the locations where the change is likely to be happened [25]. Before such stage, always a preprocessing step is necessary to produce two comparable images in the spatial domain. The process of *registration* aims at performing some geometrical operations on one of the images (or both of them), to give two compatible images, in which the pixels with the same coordinates in the two images correspond to the same physical point [28]. Many researchers have reported the impact of miss–registration on the change detection results (*e.g.*see [22]). Here, we assume that the given images are registered in the preprocess stage.

A fundamental problem in comparing two shots of the same scene in different times, is the changing recording conditions. As such, a change detection algorithms needs to be robust against atmospheric and imaging parameters. In particular, the direct solar illumination, the diffuse sky light, the path radiance, and the transmittance of the atmosphere, as well as the dark current and gain setting of the sensor may have changed individually in each spectra band [25]. All these effects can roughly be categorized into multiplicative and additive influences on the acquired data. Thus, the relation between two shots of the same scene in different times is very often modelled approximately as a linear function [21, 20]. Many authors use spectral pre–processing to compensate these changes in the imaging parameters using *ratioing* and *differentiating* approach [4].

There is a rich literature of change detection in gray–scale images. In [19] the authors survey the main approaches and compare the results. Some more frequently used methods in this field are correlation-based change detection [4], *Baysian* thresholding [5, 6], *Artificial Neural Networks* [8], and multi–block PCA [18], to name a few. Also, some researchers have tried object–based change detection, trying to locate the buildings and using morphological tools to estimate the damaged locations [3].

Compared to the single–spectrum change detection, multi–spectral change detection is being worked much less. The most frequent approach in this field is using the *principal component analysis* (PCA) to find the relation between the two images and then mark those not complying with this relation as the locations of the change [20, 25]. Some other researchers are using *regression* to find the non–changed axis [26]. Also, the application of clustering–based methods [24, 14], special color spaces [27], and the linear discriminants [7] are reported. Another idea for finding the locations of change in multispectral images is to use the exact signature of the constructing material of the scene [16]. This approach is based on expensive accurate laboratory experiments.

Here, we propose a clustering–based fuzzy change detection method for multi–spectral images and discuss its rich mathematical background. Compared to the mentioned works which result in *crisp* discrimination of changed locations, we compute a more

natural *fuzzy change detection mask* (FCDM). In fact, the approach in [25] is the special case of our proposed method when neglecting the fuzziness factor and setting the number of clusters equal to one. To formalize the comparison, in [25] global PCA [12] is used while our proposed method uses local PCA [11] to gain more adaptivity to non–linearities. Also [25] does not work in fuzzy domain.

The rest of this paper is organized as follows: Discussing the case of appropriate likelihood measures in Section 2.1 and *fuzzy principal component analysis* (FPCA) in Section 2.2 we enter to the mathematics of the problem. Then the proposed general clustering algorithm is introduced in Section 2.4 which is used for the proposed multi–spectral clustering in Section 2.5) resulting in the proposed multi–clustering fuzzy segmentation method in Section 2.6. The proposed multi–spectral fuzzy change detection method discussed in Section 2.7 uses the before mentioned methodology and gives the final result. Section 3 contains the experimental results and discussions, and finally, Section 4 concludes the paper.

2. PROPOSED METHOD

2.1. Likelihood Measurement

The *Euclidean* distance is the most generally used likelihood measure, defined as $e_r^E(\vec{c}) = (\vec{c}-\vec{\eta})^T(\vec{c}-\vec{\eta})$, where $\vec{c}$ is the vector that we intend to measure its distance to the cluster r, $\vec{x}^T$ denotes the transpose operations, and $\vec{\eta}$ is the expectation of the color vectors of cluster r defined as $\vec{\eta} = \frac{1}{n}\sum_{i=1}^{n}\vec{x}_i$.

The *Mahalanobis* distance is also a well-known likelihood measure, defining the membership of $\vec{c}$ to the cluster r as $e_r^M(\vec{c}) = (\vec{c}-\vec{\eta})^T C^{-1}(\vec{c}-\vec{\eta})$, where, C is the covariance matrix of the color vectors of the cluster r defined as $C = \frac{1}{n}\sum_{i=1}^{n}(\vec{x}_i-\vec{\eta})(\vec{x}_i-\vec{\eta})^T$.

In [1] the authors proposed to use the error made by neglecting the two least important principal components (the second and the third), as a likelihood measure. The *linear partial reconstruction error*–based (LPRE) likelihood of the vector $\vec{c}$ to the cluster r is defined as $e_r^R(\vec{c}) = \|\vec{v}^T(\vec{c}-\vec{\eta})\vec{v} - (\vec{c}-\vec{\eta})\|$, where $\vec{v}$ shows the direction of the first principal component and $\|\vec{x}\|$ denotes the normalized L_1 norm $\|\vec{x}\| = \frac{1}{N}\sum_{i=1}^{N}|x_i|$.

2.2. Fuzzy Principal Component Analysis (FPCA)

Consider, performing the PCA transform on the set of vectors $\vec{x}_i, i = 1, \cdots, N$, while the samples are members of the discrete field $\Phi = \{\vec{\phi}_i | i = 1, \cdots, n\}$. This situation happens when we work on color vectors, which are repetitions of the vectors available in $\Phi = (N_{255} \cup \{0\})^3$. Hence, we are facing the PCA problem for the set of fuzzy vectors $\{(\vec{\phi}_i; p_i) | i = 1, \cdots, n\}$, where p_i is the number of repetitions of $\vec{\Phi}_i$ (the histogram). Consider the general problem of finding the principal components of $\{\vec{x}_i; p_i | i = 1, \cdots, n\}$, when not restricting $\sum_{i=1}^{n} p_i$ to be 1. Here, as we are only concerned with the zero and one dimensional representation of the data cloud, we will derive the formulation for $\vec{\eta}$ and $\vec{\omega}_1$, but the same method leads to the computation of the other fuzzy principal components as well.

All objective functions in the PCA theory are of the form,

$$\Delta(\vec{x}_\circ) = \sum_{i=1}^{n} Q_{\vec{x}_\circ}(\vec{x}_i), \tag{1}$$

where $Q_{\vec{x}_\circ}(\vec{x}_i)$ is a quadratic function. Assuming the case of repeated vectors, the objective function changes to,

$$\Delta(\vec{x}_\circ) = \sum_{i=1}^{n} p_i Q_{\vec{x}_\circ}(\vec{\phi}_i), \tag{2}$$

where, p_i is the number of occurrence of $\vec{\Phi}_o$. Thus, it is reasonable to define the objective functions in the fuzzy domain as,

$$\Delta(\vec{x}_\circ) = \sum_{i=1}^{n} p_i Q_{\vec{x}_\circ}(\vec{x}_i). \tag{3}$$

Note that the case of $p_i \equiv 1$ leads to the same non–fuzzy definition. Also the assumption of $\sum_{i=1}^{n} p_i = 1$ is not necessary, because it shows itself as a constant scaling factor; not affecting the minima.

The objective function for the expectation vector in the crisp domain is defined as $Q_{\vec{x}_\circ}(\vec{x}) = \|\vec{x} - \vec{x}_\circ\|^2$ [10]. Thus, the fuzzy expectation is the minima of the objective function defined as $\Delta_\circ(\vec{x}_\circ) = \sum_{i=1}^{n} p_i\|\vec{x}-\vec{x}_\circ\|^2$. Investigating the condition in which the derivative of $\Delta_\circ(\vec{x}_\circ)$ in terms of $\vec{x}_\circ$ gets zero, yields,

$$\vec{\eta} = \frac{\sum_{i=1}^{n} p_i \vec{x}_i}{\sum_{i=1}^{n} p_i}. \tag{4}$$

The objective function for the crisp first principal direction is the direction of the maximum deviation or equivalently the direction of minimum one–dimensional reconstruction error, defined as [10],

$$Q_{\vec{x}_\circ}(\vec{x}) = \|\vec{x} - \vec{\eta} - \vec{x}_\circ^T(\vec{x} - \vec{\eta})\vec{x}_\circ\|^2. \tag{5}$$

Thus, in the fuzzy domain, we should minimize the objective function defined as,

$$\Delta_I(\vec{x}_\circ) = \sum_{i=1}^{n} p_i\|\vec{x} - \vec{\eta} - \vec{x}_\circ^T(\vec{x} - \vec{\eta})\vec{x}_\circ\|^2. \tag{6}$$

Algebraic derivation of (6) and incorporating the size constraint on the principal components, ($\|\vec{x}_\circ\|^2 = 1$), leads to,

$$\Delta_I(\vec{x}_\circ) = \sum_{i=1}^{n} p_i\|\vec{x}_i - \vec{\eta}\|^2 - \vec{x}_\circ^T\tilde{C}\vec{x}_\circ, \tag{7}$$

where, $\tilde{C}$ is the *fuzzy covariance matrix* defined as,

$$\tilde{C} = \sum_{i=1}^{n} p_i(\vec{x}_i - \vec{\eta})(\vec{x}_i - \vec{\eta})^T. \tag{8}$$

As the first term in (7) is not a function of $\vec{x}_\circ$, thus, $\Delta_I(\vec{x}_\circ)$ is minimized when the second term is maximized. Using the method of *Lagrange* multipliers for embedding the $\|\vec{x}_\circ\| = 1$ constraint, we should maximize the objective function defined as,

$$\bar{\Delta}_I(\vec{x}_\circ) = \vec{x}_\circ^T\tilde{C}\vec{x}_\circ + \lambda(\|\vec{x}_\circ\|^2 - 1). \tag{9}$$

Differentiating (9) in terms of $\vec{x}_\circ$ and assigning the result to zero we get $\tilde{C}\vec{x}_\circ = \lambda\vec{x}_\circ$. Thus, $\vec{x}_\circ$ is an eigenvector of $\tilde{C}$, resulting in $\tilde{\Delta}_I(\vec{x}_\circ) = \lambda$. As a result, the direction of the first fuzzy principal component is the eigenvector of $\tilde{C}$, corresponding to the largest eigenvalue. It can be easily proved (using the same method), that the other principal directions correspond to the other eigenvectors of $\tilde{C}$, sorted by the eigenvalues in a descending fashion; like the classic PCA.

2.3. Weighted Powered Sum Minimization

Consider the optimization problem defined as minimizing the function,

$$\Delta(x_1, \cdots, x_n) = \sum_{i=1}^{n} x_i^m w_i, \tag{10}$$

under the assumption of $\sum_{i=1}^{n} x_i = 1$. Substituting $x_p = 1 - \sum_{i=1, i\neq p}^{n} x_i$ in (10) and differentiating with respect to x_i for $i \neq p$ results in,

$$x_p = \frac{w_p^{-\frac{1}{m-1}}}{\sum_{i=1}^{n} w_i^{-\frac{1}{m-1}}}. \tag{11}$$

Note that, hereby the normality constraint is satisfied. Substituting (11) in (10) we have,

$$\Delta = \left(\sum_{i=1}^{n} w_i^{-\frac{1}{m-1}} \right)^{-(m-1)}. \tag{12}$$

Assuming $m > 1$, Δ will be less than all w_p. Also, note that the case of setting all x_i equal to zero, except for $x_p = 1$, leads to $\Delta = w_p$. Hence, having in mind that the gradient of Δ gets zero just once, while the value of Δ at that point is smaller than some marginal points, thus, (11) leads to the global minima of the problem in hand.

2.4. Proposed General Clustering Algorithm

Consider the general clustering problem stated as minimizing the objective function defined as,

$$J(X, \Phi) = \sum_{i=1}^{n} \sum_{j=1}^{c} p_{ij}^m D_{ij}. \tag{13}$$

This describes the best choice of clustering the data points $X = \{\vec{x}_1, \cdots, \vec{x}_n\}$ into c clusters described by $\Phi = \{\phi_1, \cdots, \phi_c\}$. Here, p_{ij} is the fuzzy membership of $\vec{x}_i$ to the jth cluster and D_{ij} is the distance between this point and the cluster. Assume that,

$$D_{ij} = \Psi(\vec{x}_i, \phi_j), \tag{14}$$

is the *appropriate* distance function for the vector geometry under investigation. Consequently, under the Ψ distance function, we have $p_{ij} \propto D_{ij}$ and $p_{ij} \propto D_{pj}^{-1}$ for $p = 1, \cdots, n, p \neq i$. Here, ϕ_j is the defining parameters of the jth cluster according to the general cluster model.

As an special case, the FCM models each cluster with a single vector ($\phi_j = \{\vec{\eta}_j\}$), and defines the Ψ function as the squared *Euclidean* distance between the given vector and the cluster center. The key point of the proposed *general fuzzy clustering* (GFC) algorithm is the deep relation between the cluster model and the cluster tuning function. Note that for a set of given points, the expectation vector minimizes the sum of the squared *Euclidean* distances. Now assume that the function Υ tunes the cluster model ϕ_j to best fit the points. This means that for the fuzzy set $\tilde{X} = \{(\vec{x}_i; p_i) | i = 1, \cdots, n\}$ of vectors, $\phi = \Upsilon(\tilde{X})$ is the solution for;

$$\phi^* = \arg_\phi \min \left\{ \sum_{(\vec{x}_i, p_i) \in \tilde{X}} p_i \Psi(\vec{x}_i, \phi) \right\}. \tag{15}$$

Then, the function $\Upsilon(\tilde{X}) \equiv \tilde{E}\{\tilde{X}\}$ is the solution for $\Psi(\vec{x}, \vec{\eta}) = \|\vec{x} - \vec{\eta}\|^2$. Here, $\tilde{E}\{\tilde{X}\}$ stands for the fuzzy expectation of a fuzzy set.

Back to the main problem of minimizing (13), assume that we have the dual functions, $\Psi(\cdot)$ and $\Upsilon(\cdot)$. Here, we propose an algorithm that converges to a minimal point of (13), if at least one exists. Consider rewriting (13) as,

$$J(X, \Phi) = \sum_{i=1}^{n} \Delta_i, \Delta_i = \sum_{j=1}^{c} p_{ij}^m D_{ij}. \tag{16}$$

Also, assume that we have fixed D_{ij} and we are trying to decline $J(X, \Phi)$ by working on p_{ij}. As the only restriction on p_{ij} is the normality condition ($\forall i, \sum_{j=1}^{c} p_{ij} = 1$), there is no connection between p_{ij} and $p_{i'j}$, for $i \neq i'$. Thus, declining Δ_is independently results in declination of $J(X, \Phi)$. Having in mind the normality constraints, minimizing Δ_i is a weighted powered sum minimization (discussed in Section 2.3). Hence, we get,

$$p_{ij} = \frac{D_{ij}^{-\frac{1}{m-1}}}{\sum_{k=1}^{c} D_{ik}^{-\frac{1}{m-1}}}. \tag{17}$$

Note that during this stage, $J(X, \Phi)$ is declined, except for the case that p_{ij}s satisfy (17) at first. We will come back to this situation later. Now, consider rewriting (13) as,

$$J(X, \Phi) = \sum_{j=1}^{c} \Theta_j, \Theta_j = \sum_{i=1}^{n} p_{ij}^m D_{ij}. \tag{18}$$

Assume that we have fixed p_{ij} and we are trying to decline $J(X, \Phi)$ by working on D_{ij}. Declining Θ_j means tuning the j–th cluster to minimize the overall distances in a fuzzy scheme. Note that like above, it can be solved independently for different clusters, resulting in overall decline of the $J(X, \Phi)$. The solution in this state is obtained by using the Υ function as,

$$\phi_j = \Upsilon\left(\left\{ (\vec{x}_i, p_{ij}^m) \middle| i = 1, \cdots, n \right\} \right). \tag{19}$$

Again, the $J(X, \Phi)$ can never increase in this state, and its constancy shows that all the clusters have been in the best place, according to the special distance function Ψ. The halting test of the algorithm is the stationarity of the cluster parameters measured by the average cluster center change defined as,

$$\delta = \sqrt{\frac{1}{c} \sum_{i=1}^{c} \|\vec{\eta}_i - \vec{\eta}_{i'}\|^2}, \tag{20}$$

getting smaller than a preselected threshold δ_T. We propose the procedure shown in Figure 1 to cluster the given data.

Note that during the two stages, $J(X, \Phi)$ never rises and the case of it being constant in two consecutive stages results in staying constant for all coming stages. Thus, using this method, $J(X, \Phi)$ is going towards a minimum point, and never gets oscillated. Note that rather than the FCM, other methods like GK, FEC, and FCV, are also special cases of the proposed general clustering method.

Aim: Clustering data.
Inputs:
Input vectors $\vec{x}_i, i = 1, \cdots, n$.
Number of the clusters c.
Distance function Ψ.
Stationarity threshold δ_T.
Output:
A set of c clusters $\phi_i, i = 1, \cdots, c$.
Method:
[1] Randomly initialize clusters ϕ_i, satisfying the model constraints.
[2] Compute D_{ij}, using (14).
[3] Compute p_{ij}, using (17).
[4] Store the clusters as $\phi'_i = \phi_i$.
[5] Update clusters, using (19).
[6] Compute δ, using (20).
[7] If $\delta > \delta_T$ goto 2, else
[8] return $\phi_i, i = 1, \cdots, c$.

Figure 1: Proposed clustering procedure.

2.5. Proposed Multi–Spectral Clustering

When attempting to segment an image, the first step is to find a few classes that define the image content in an intuitive way. As stated in Section 2.1, the LPRE distance defined as,

$$\Psi(\vec{c}, [\vec{\eta}, \vec{v}]) = \|(\vec{c} - \vec{\eta}) - \vec{v}^T(\vec{c} - \vec{\eta})\vec{v}\|^2, \tag{21}$$

results in a good subjective clustering of the spectral vectors in color fields. In this approach, the cluster model is a *cylinder* with its central axis having $\vec{\eta}$ and parallel with $\vec{v}$ (see Figure 2). Also, the dual function $\Upsilon(\tilde{X})$ computes the fuzzy expectation and the first fuzzy principal component of $\tilde{X}$ as new values of $\vec{\eta}$ and $\vec{v}$ (as discussed in Section 2.2). It must be emphasized that this formulation results in an special case of the FCV method ($r = 1$) [13].

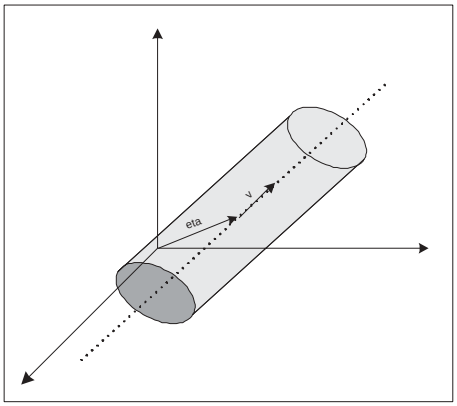


Figure 2: Cylindrical cluster model.

2.6. Proposed Multi–Spectral Fuzzy Segmentation

We face the segmentation problem as applying the maximum likelihood on the cluster membership information to relabel the image pixels in the way that it results in meaningful geometrical continuous segments. Consider, clustering an image I, into c clusters by using the method proposed in Section 2.5. The result is the set of c images $J_i, i = 1, \cdots, c$, which shows the likelihood of each pixel to each of the c clusters. Note that J_i satisfies, $\forall x, y : \sum_{i=1}^{n} J_i(x, y) = 1$. Now, consider a $p \times p$ smoothing convolution kernel M. Note that M satisfies $\sum_{i=1}^{p}\sum_{i=1}^{p} M_{ij} = 1$. Hence, applying M to each J_i independently to acquire $\tilde{J}_i$, the new membership maps will also satisfy the normality condition. Thus, $\tilde{J}_i, i = 1, \cdots, c$ can be considered as the smoothed likelihood to the c clusters. The main benefit of using $\tilde{J}_i$ over J_i is the smoother resulting segments. The crisp segmentation results can also be obtained using the ML. Here, we propose to use an averaging kernel as M.

2.7. Proposed Multi–Spectral Fuzzy Change Detection

Assume the two images I_1 and I_2 which are registered using some processing method. Also, assume that I_1 is segmented into c classes of $\phi_i, i = 1, \cdots, c$, where $J_{ixy}, i = 1, \cdots, c$ shows the membership of $\vec{I}_{1xy}$ to the i–th class. Assume performing the FPCA on the fuzzy set $\{(\vec{I}_{2xy}; J^m_{ixy})\}$ to find the new clusters $\tilde{\phi}_i, i = 1, \cdots, c$ and computing the new membership values $\tilde{J}_{ixy}$, which show the degree of membership of $\vec{I}_{2xy}$ to the i–th new class $\tilde{\phi}_i$. We propose,

$$\delta_{xy} = \sqrt{\frac{1}{c}\sum_{i=1}^{c}(J_{ixy} - \tilde{J}_{ixy})^2}, \tag{22}$$

as the probability of the point (x, y) being changed.

Note that if $I_1 \equiv I_2$, J_{ixy} and $\tilde{J}_{ixy}$ will be identical, resulting in $\delta_{xy} \equiv 0$, as desired. Now assume that there is no change between the two images I_1 and I_2, unless for the changes in the imaging conditions. Assume that $\vec{x}_i$ and $\vec{y}_i$ are the spectral vectors of the same physical points in the two images I_1 and I_2, respectively. As stated in Section 1, $\vec{x}_i$ and $\vec{y}_i$ relate through a linear transform namely $\vec{x}_i = A\vec{y}_i + \vec{b}$. Here, we model A as a non–singular invertible matrix with its eigenvalues being almost constant. The matrix A in the *singular value decomposition* (SVD) form is written as $A = VDU^{-1}$, where, U and V are orthogonal matrices and D is a diagonal matrix with the eigenvalues of A as its elements.

The expectation vectors in the two images I_1 and I_2 relate as,

$$E\Big\{\vec{x}_i\Big\} = E\Big\{A\vec{y}_i + \vec{b}\Big\} = AE\Big\{\vec{y}_i\Big\} + \vec{b}. \tag{23}$$

The fuzzy covariance matrices of the two images I_1 and I_2 satisfy the following equations,

$$C_1 = AE\Bigg\{\Big(\vec{y}_i - E\{\vec{y}_j\}\Big)\Big(\vec{y}_i - E\{\vec{y}_j\}\Big)^T\Bigg\}A^T = AC_2A^T \tag{24}$$

Assume that the eigenvectors of C_1 are $\vec{v}_i, i = 1, \cdots, m$ corresponding to the eigenvalues of $\lambda_i, i = 1, \cdots, m$ and the eigenvectors of C_2 are $\vec{u}_i, i = 1, \cdots, m$ corresponding to the eigenvalues of $\rho_i, i = 1, \cdots, m$. Also, assume the eigenvectors of A to be $\vec{w}_i, i = 1, \cdots, m$ corresponding to the eigenvalues of $\varepsilon_i, i = 1, \cdots, m$. Thus,

$$\forall i, C_1\vec{v}_i = \lambda_i\vec{v}_i, C_2\vec{u}_i = \rho_i\vec{u}_i, A\vec{w}_i = \varepsilon_i\vec{w}_i. \tag{25}$$

First assume that the eigenvectors of A are all exactly equal to the fixed value of λ (or equivalently $\forall i, \varepsilon_i = \lambda$). Thus,

$$A = VDU^{-1} = Vdiag(\lambda, \cdots, \lambda)U^{-1} = \lambda VU^{-1}. \quad (26)$$

In this situation,

$$A^T = \lambda UV^{-1} = \lambda^2 A^{-1} \rightarrow A^T A = AA^T = \lambda^2 I. \quad (27)$$

Now, note that,

$$C_1 A\vec{u}_i = AC_2 A^T A\vec{u}_i = \lambda^2 AC_2\vec{u}_i = \lambda^2 \rho_i A\vec{u}_i. \quad (28)$$

Thus, $A\vec{u}_i$ is the eigenvector of C_2 corresponding to the eigenvalue of $\lambda^2\rho_i$. Note that,

$$\|A\vec{u}_i\| = \lambda\|\vec{u}_i\| = \lambda. \quad (29)$$

As the eigenvalues and eigenvectors of a single matrix are identical, we have,

$$\Big\{(\frac{1}{\lambda}A\vec{u}_1, \lambda^2\rho_1), \cdots, (\frac{1}{\lambda}A\vec{u}_m, \lambda^2\rho_m)\Big\} = \Big\{(\vec{v}_1, \lambda_1), \cdots, (\vec{v}_m, \lambda_m)\Big\}. \quad (30)$$

As $\lambda^2 > 0$ we have,

$$\forall i, \vec{v}_i = \frac{1}{\lambda}A\vec{u}_i, \lambda_i = \lambda^2\rho_i. \quad (31)$$

Thus, using the above re–clustering method, the cluster $\phi = [\vec{\eta}, \vec{v}]$ in I_2 results in the cluster $\tilde{\phi} = [A\vec{\eta} + \vec{b}, A\vec{v}]$. Now, we have,

$$\Psi(\vec{x}_i, \tilde{\phi}) = \Big\|\Big[(A\vec{y}_i + \vec{b}) - (A\vec{\eta} + \vec{b})\Big] - \frac{1}{\lambda^2}\vec{v}^T A^T \Big[(A\vec{y}_i + \vec{b}) - (A\vec{\eta} + \vec{b})\Big] A\vec{v}\Big\|^2 \quad (32)$$

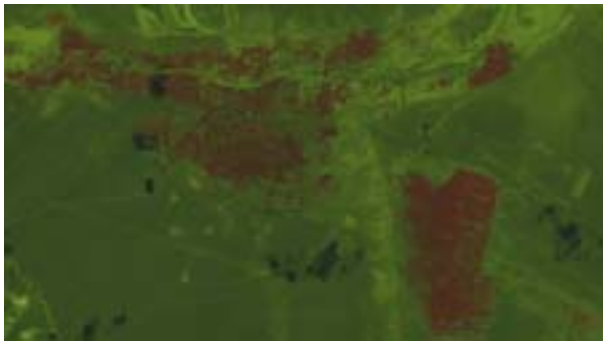
$$\Psi(\vec{x}_i, \tilde{\phi}) = \Big\|A(\vec{y}_i - \vec{\eta}) - \vec{v}^T(\vec{y}_i - \vec{\eta})A\vec{v}\Big\|^2 = \lambda^2\Psi(\vec{x}_i, \tilde{\phi}) \quad (33)$$

$$\tilde{J}_{ixy} = \frac{\Psi(\vec{x}_i, \tilde{\phi}_j)^{-\frac{1}{m-1}}}{\sum_{k=1}^{c}\Psi(\vec{x}_i, \tilde{\phi}_k)^{-\frac{1}{m-1}}} = \frac{\Psi(\vec{x}_i, \phi_j)^{-\frac{1}{m-1}}}{\sum_{k=1}^{c}\Psi(\vec{x}_i, \phi_k)^{-\frac{1}{m-1}}} = J_{ixy}, \quad (34)$$

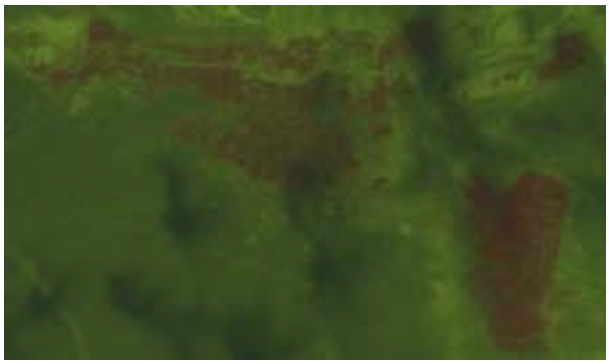
resulting in $\delta_{xy} = 0$. Thus, the proposed method will be independent of the lighting and imaging conditions.

Now, assume a more realistic case that ε_is are not exactly the same but we have $\lambda - \delta\lambda \leq \varepsilon_i \leq \lambda + \delta\lambda$. For the cases that $\frac{\delta\lambda}{\lambda}$ is too small the above equations change to semi–equations and still marginally hold. In this situation $\delta_{xy} \simeq 0$.

In contrast, physical changes result in different material in a single point in different shots. Hence, they produce absolutely different values of J_{ixy} and $\tilde{J}_{ixy}$, resulting in non–zero patterns of δ_{xy}. In the proposed method, at the same time both the image sequence segmentation and the fuzzy change detection are performed.



(a)



(b)

Figure 3: Multi–spectral images of the city of *Bam*. (a) 2003-12-04. (b) 2003-12-29. Courtesy of IRSC.

3. EXPERIMENTAL RESULTS

The tests were performed using a PIV 2600MHz personal computer with 512MB of RAM. Figure 3 shows two images of the city of *Bam* by the *IRS–1D* satellite *LIS III* sensor, each containing 3 channels.

Figure 4 shows three query regions in the image shown in Figure 3–(a), corresponding to the *urban area*, *desert*, and *river bed*, respectively. Figure 5 shows the results of segmenting the original image into three segments using the *maximum likelihood* (ML) method with the three likelihood measures of *Euclidean*, *Mahalanobis*, and LPRE. Note the cluster shapes in Figure 5–(a),(c), and (e) and the corresponding segmentation results in Figure 5–(b),(d), and (f), respectively. As Figure 5 shows the *Euclidean* and the *Mahalanobis* distances have misclassified many portions of the *desert* into the *urban* class, while the LPRE shows the best compliance with the physical classes. The same results are observed in other numerous examples. Thus, the LPRE–based likelihood measure is the best of all descriptor for the images under investigation. Specially note that unless for the LPRE, the two other measures have failed to classify the non–urban areas in numerous situations and have included them in the urban class. This is so important, because after the earthquake, some spectral vectors in the urban class are getting members of the non–urban class, due to the destruction.

Figure 6 shows the result of applying the proposed multispectral segmentation method on the image shown in Figure 3–(a), with the parameters set as, $c = 3$, $m = 1\frac{1}{3}$, $\delta_T = \frac{1}{2}$, and $k = 2$.

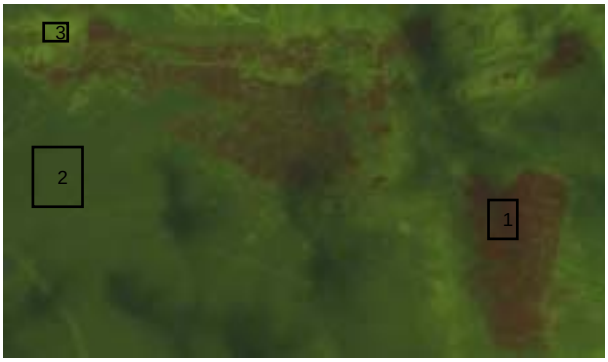


Figure 4: Three query regions in the image shown in Figure 3–(a).

Figure 7 shows the result of the proposed multi–spectral fuzzy change detection applied on the image shown in Figure 3 with the parameters set as $c = 3$, $m = 1\frac{1}{3}$, $\delta_T = \frac{1}{2}$, and $k = 2$. To get a better printing result, the values of $1 - \delta_{xy}$ are shown here. Note the rightmost–bottom part of the city, were destructed parts are visible, and the corresponding black spots in the change map. Also, note that the desert part of the image is ranked to be less changed than the urban part, as desired. The main drawback of the proposed method is its sensitivity to the clouds. We propose using the channels which are less sensitive to the spectrum of the clouds.

Table 1 compares the proposed method with the available literature. The comparison is made in terms of the dimension of the input data (multispectral vs. grayscale), model–based tendency of the method, locality of change detection (adaptability to different effects of the change in different spectral bands), spatial information inclusion, human intervention (supervised vs. unsupervised), and the output space (crisp vs. fuzzy). Model–based methods are those using the prior knowledge about the image, such as information about the scale, contents, material and so on. As shown in Table 1, while the proposed method is the only fuzzy–output one, it works unsupervised and processes the changes locally with no need to prior information about the image contents.

4. CONCLUSIONS

The performance of three likelihood measures are investigated, in terms of their efficiency in segmenting multi–spectral images and the LPRE distance is shown to outperform the *Euclidean* and the *Mahalanobis* measures. A new fuzzy clustering method is proposed which uses the proposed likelihood measure, and its efficiency is shown both theoretically and experimentally. A new fuzzy change detection method is proposed which measures the changes in the membership values of the spectral vectors before and after the change. The method is mathematically proved to be efficient while experiments confirm the results. Comparison of the proposed method with the available literature proves it to be superior in terms of giving fuzzy results, working unsupervised, processing the changes locally, and not needing prior information about the image contents.

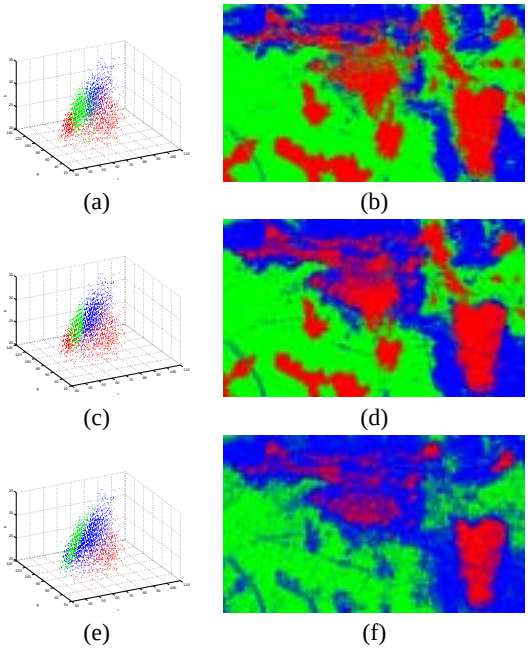


Figure 5: The results of applying maximum likelihood on the query regions shown in Figure 4. (a) *Euclidean*. (b)*Mahalanobis*. (c) LPRE.

Acknowledgement

We would like to appreciate the valuable discussions and suggestions made by professor *Y. Kosugi* and professor *M. Nakamura* from *Tokyo Institute of technology*. We also wish to thank the *Iranian Remote Sensing Center* (IRSC) for providing us with the remote sensing images used in this paper. The first author also wishes to thank Ms. *Azadeh Yadollahi* for her encouragement and invaluable ideas.

5. REFERENCES

[1] A. Abadpour and S. Kasaei, "A new parametric linear adaptive color space and its pca–based implementation," in *The 9th Annual CSI Computer Conference, CSICC*, Tehran, Iran, 2004, pp. 125–132.

[2] J. R. Anderson, E. E. Hardy, J. T. Roach, and R. E. Witmer, "A land use and land cover classification system for use with remote sensor data," United States Government Printing Office, Washington, 1976.

[3] G. Andre, L. Chiroiu, C. Mering, and F. Chopin, "Building destruction and damage assessment after earthquake using high resolution optical sensors. the case of the gujarat earthquake of 26, 2001," Unknown IEEE Conference., 2003.

[4] K. M. Bergen, D. G. Brown, J. R. Rutherford, and E. J. Gustafson, "Development of a method for remote sensing of land–cover change 1980–200 in the ufs north central region using heterogeneous usgs luda and noaa avhrr 1km data," in *Proceedings, International Geoscience and Remote Sensing Symposium*, Toronto, CA, 2002.

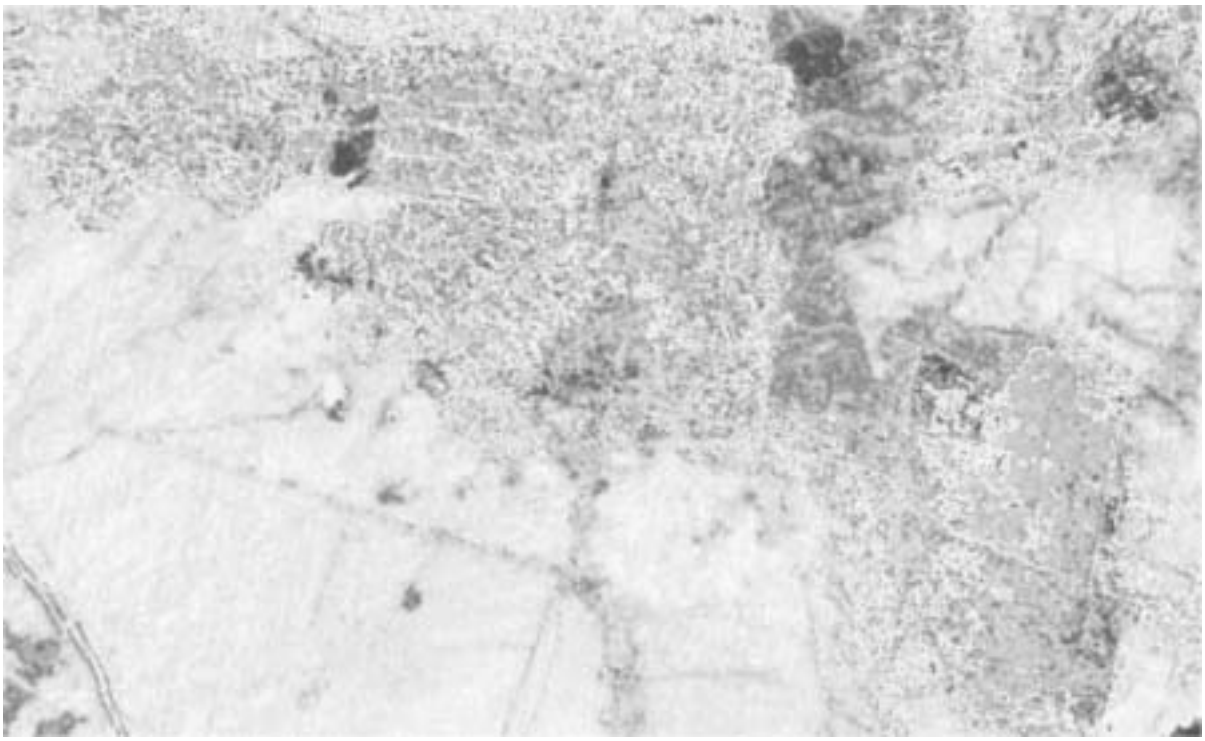


Figure 7: Result of the proposed multi–spectral fuzzy change detection method applied on the images shown in Figure 3.

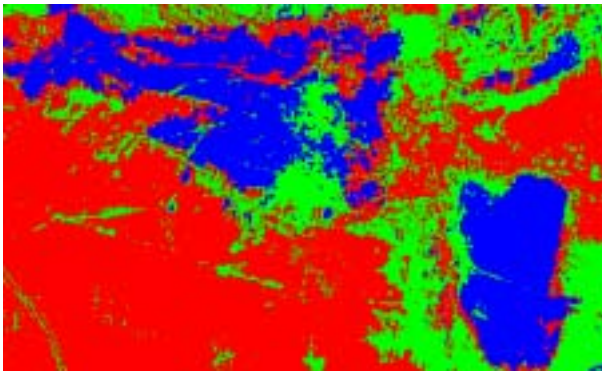


Figure 6: Result of the proposed segmentation method applied on the image shown in Figure 3–(a).

[5] L. Bruuzzone and D. F. Preito, "A bayesian approach to automatic change detection," in *Proceedings of the IEEE 1999 Int. Geoscience and Remote Sensing Symposium (IGARSS99)*, vol. 3, Hamburg, Germany, 1999, pp. 1816–1818.

[6] L. Bruzzone and R. Cossu, "Analysis of multitemporal remote–sensing images for change detection: Bayesian thresholding approaches," University of Trento, Department of Information and Communication Technology, 38050 Povo- Trento (Italy), Tech. Rep. DIT-02-0031, 2003.

[7] M. J. Canty, I. Niemeyer, and Q. B. Truong, "Change detection with multispectral satellite imagery: Application to nuclear verification," in *Proceedings of the International Symposium on Spectral Sensing Research, ISSSR 2001 (CD-Rom)*, Quebec City, 2001.

[8] X. Dai and S. Khorram, "Development of a new automated land cover change detection system from remotely sensed imagery based on artificial neural networks," in *Proceedings of 1997 IEEE International Geoscience and Remote Sensing Symposiums (IGARSS'97)*, Singapore, 1997, pp. 1029–1031.

[9] C. S. Fischer and L. M. Levinen, "Monitoring california's hardwood rangelands using remotely sensed data," in *Proceedings of the Fifth Oak Symposium*, Oak Woodlands, 2001.

[10] D. Gering, "Linear and nonlinear data dimensionality reduction," MIT Area Exam, http://www.ai.mit.edu/people/ gering/areaexam/, 2002.

[11] K. Honda, N. Sugiura, and H. Ichihashi, "Robust local principal component analyzer with fuzzy clustering," in *IJCNN 2003 Conference Proceedings*, 2003, pp. 732–737.

[12] A. Hyvrinen and E. Oja, "Independent component analysis: Algorithms and applications," *Neural Networks*, vol. 13(4–5), pp. 411–430, 2000.

[13] J. M. Leski, "Fuzzy c–varieties/elliptotypes clustering in reproducing kernel hilbert space," *Fuzzy Sets and Systems*, vol. 141, pp. 259–280, 2004.

[14] A. M. Massone, F. Masulli, and A. Petrosini, "Fuzzy clustering algorithms and landsat images for detection of waste

Table 1: Comparison of the proposed method with the literature.

Method	**Multi–spectral**	**Non Model–based**	**Local**	**Spatial**	**Unsupervised**	**Fuzzy**
Yamamoto et. al.[26]	√	√	–	√	–	–
Kathleen et. al.[4]	–	√	–	–	–	–
Nakamura et. al.[17]	–	–	–	√	–	–
Matsuoka et. al.[16]	–	√	–	–	–	–
Bruzzone et. al.[6]	–	√	–	–	–	–
Dai et. al.[8]	–	–	–	–	–	–
Canty et. al.[7]	√	√	–	√	–	–
Wiemker et. al.[25]	√	√	–	–	√	–
Qiu et. al.[18]	–	√	√	√	–	–
Mitomi et. al.[16]	√	–	–	–	–	–
Andre et. al.[3]	–	–	–	–	–	–
Proposed Method	√	√	√	√	√	√

areas: A comparison," *Advances in Fuzzy Systems and Intelligent Technologies*, pp. 165–175, 2000.

[15] M. Matsuoka and F. Yamazaki, "Application of the damage detection method using sar intensity images to recent earthquakes," in *Proceedings of the International Geoscience and Remote Sensing Symposium, IEEE,*, 2002, pp. CD–ROM.

[16] H. Mitomi, M. Matsuoka, F. Yamazaki, H. Taniguchi, and Y. Ogawa, "Determination of the areas with building damage due to the 1995 kobe earthquake using airborne mss images," in *IEEE International Geoscience and Remote Sensing Symposium, CD–ROM*, 2002.

[17] M. Nakamura, M. Sakamoto, S. Kakumoto, and Y. Kosugi, "Stabilizing the accuracy of change detection from geographic images by multi–levelled exploration and selective smoothing," in *Proceedings of GIS2003*, Vancouver, 2003.

[18] B. Qiu, V. Prinet, E.Perrier, and O. Monga, "Multi–block pca method for image change detection," in *12th International Conference on Image Analysis and Processing (ICIAP'03)*, Mantova, Italy, 2003, p. 385.

[19] R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: A systematic survey," *IEEE Transactions on Image Processing (Submitted)*, 2004.

[20] J. A. Richards, *Remote Sensing Digital Image Analysis*. Heidelberg, New york: Springer, 1993.

[21] A. Singh, "Review article: Digital change detection techniques using remotely–sensed data," *International Journal of Remote Sensing*, vol. 10(6), pp. 989–1003, 1989.

[22] J. R. G. Townshend, C. O. Justice, and C. Gurney, "The impact of misregisitration on change detection," *IEEE Transaction on Geoscience and Remote Sensing*, vol. 30(5), pp. 1054–1060, 1992.

[23] P. S. Tradile and R. C. Conglaton, "A change–detection analysis: Using remotely sensed data to assess the progression of development in essex county, massachusetts from 1999 to 2001," http://www.unh.edu/natural–resources/students.html.

[24] M. Trivedi and J. Bezdek, "Low–level segmentation of aerial images with fuzzy clustering," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 16(4), p. 589598, 1986.

[25] R. Wiemker, A. Spek, D. Kulbach, H. Spitzer, and H. Bienlein, "Unsupervised robust change detection on multispectral imagery using spectral and spatial features," in *Proceedings of the Third International Airborne Remote Sensing Conference and Exhibition*, Copenhagen, Denmark, 1997.

[26] T. Yamamoto, H. Hanazumi, and S. Chino, "A change detection method for remotely sensed multi–spectral and multi–temporal images using 3–d segmentation," *IEEE Transaction on GRS*, vol. 39(5), pp. 976–985, 2001.

[27] Y. Yanamura and H. Saji, "Automatic registration of aerial image and digital map for detection of earthquake damaged area," in *Proceedings of the VIIth Digital Image Computing: Techniques and Applications*, Sydney, 2003, pp. 117–126.

[28] B. Zitova and J. Flusser, "Image registration methods: A survey," *Image and Vision Computing*, vol. 21, pp. 977–1000, 2003.