

Change Vector Analysis using Enhanced PCA and Inverse Triangular Function-based Thresholding

Munmun Baisantry*, D.S. Negi, and O.P. Manocha

Defence Electronics Applications Laboratory, Dehradun-248 001, India

**E-mail: munmunbaisantry@gmail.com*

ABSTRACT

Change vector analysis is a very sophisticated method to evaluate land-use/land-cover changes meaningfully. By making proper choice of input data in the form of bands (for instance, red, NIR etc) or features (for instance, greenness, brightness, wetness etc), information about both the magnitude as well as the type/nature of changes can be extracted. However, improper selection of thresholds is always a hindrance to a good change detection algorithm. The paper has proposed an improved technique to select threshold appropriately by means of principal component difference and inverse triangular function. The changes have been represented using class-based circular wheel representation. Results have been shown to further testify the performance of proposed algorithm.

Keywords: Change vector analysis, principal component difference, inverse triangular function, Kauth-Thomas transformation

$$r = \sqrt{((\Delta A_i)^2 + (\Delta B_i)^2)} \quad \theta = \cos^{-1} \left(\frac{(\Delta A_i)}{r} \right),$$

$$\Phi = \cos^{-1} \left(\frac{(\Delta B_i)}{r} \right) \tag{1}$$

1. INTRODUCTION

This paper presents a systematic approach to detect and describe different types of changes in satellite imagery using change vector analysis (CVA) and a unified change/no-change thresholding method using principal component difference (PCD) and inverse triangular function. The algorithm begins with geometric registration which is the first pre-processing stage. This was followed by radiometric normalization to remove the false changes due to atmospheric differences in two temporal images. Change/no-change information as well as the degree of change between the geometrically registered and radiometrically normalized images was acquired by means of the thresholding technique using PCD, inverse triangular function. Finally, the type of change on the changed pixels was determined using CVA and are represented by circular wheel representing scheme.

Commonly used change detection techniques like image differencing, image ratio, etc. use only one band for retrieving change/no-change information. The disadvantage associated is that information conveyed by the other bands is not made use of. Further, these methods cannot identify the types of changes that have occurred. Although, the first problem can be taken care of by principle component analysis (PCA) which combines essential information from all the bands in the first few components, it still cannot determine the type /nature of occurred changes. The potential of CVA lies with the fact that it can give meaningful information about the type of change¹⁻⁴. The information, still, however depends upon the bands chosen in the algorithm. The magnitude and direction of change can be characterized in Eqn (1).

$$r = \sqrt{((\Delta A_i)^2 + (\Delta B_i)^2)} \quad \theta = \cos^{-1} \left(\frac{(\Delta A_i)}{r} \right),$$

where r , θ , Φ are the magnitude and directions respectively. Instead of spectral bands, one can also use combinations of features like greenness, brightness, normalized differential vegetation index, brightness index, etc.

Principal component analysis is a method that reduces data dimensionality and is suitable for data sets in multiple dimensions, for instance combining information from all bands of multispectral satellite imagery. Principal component analysis seeks for directions where maximum information variance is possible yet the information conveyed along one direction should not be replicated/ correlated in/to the other directions. Given a cloud of data, the first PCA component would be along the midline; the next component would be orthogonal to the first one, so on and so forth^{5,6}.

2. DATASET USED AND STUDY AREA

The image dataset used in this study is obtained from IRS-1D satellite. The satellite has a multispectral sensor LISS-III operating in four spectral bands, two visible, one near infrared (spatial resolution 24.5 m) and one SWIR band (spatial resolution 70 m) as well as a PAN sensor (spatial resolution 5.8 m).

The dataset comprises of two sets of geometrically registered multispectral images of same area taken on two different dates. The reference and subject images are cloud free and also have a spectrally good dynamic range. For CVA using red-NIR bands, FCC composite images acquired with green, red and NIR bands have been used as shown in Figs 3(a) and 3(b). For CVA using greenness-brightness, all the four bands

have been used. FCC composite images acquired with green, red and NIR bands are shown in Figs 4(a) and 4(b).

3. METHODOLOGY

The whole change detection process can be divided into six steps: (i) pre-processing (radiometric normalization using MAD transformation), (ii) change criteria function, (iii) contribution of PCDs, (iv) determination of inverse triangular function, (v) degree of change, and (vi) change representation methodology.

3.1 Radiometric Normalization (Multivariate Alteration Detection)

The multi-spectral images are geometrically registered to each other. The images will be made radiometric similar to each other to remove any changes arising between them due to atmospheric and calibration differences. For this, automatic radiometric normalization technique is developed which is based on multivariate alteration detection (MAD)^{7,8}. Pseudo invariant features (PIFs) are selected automatically using MAD transformation and the normalization coefficients are calculated using robust linear regression techniques (Iteratively reweighted least squares). Selection of PIFs is based upon the assumption that atmospheric and calibration differences are linearly related and MAD components are invariant to such affine transformations. This method is based on canonical correlation analysis (CCA)⁹. Canonical correlation analysis is basically a method to find a set of basis vectors for two sets of multidimensional variables such that their linear relationship is maximally correlated. Canonical correlation analysis seeks a pair of linear transformations one for each of the sets of variables such that when the set of variables are transformed the corresponding coordinates are maximally correlated.

Assuming that we have two temporal images having *N* spectral channels, we can represent intensities of both images by random vectors *F* and *G*, where $F = [F_1, F_2, F_3, \dots, F_N]^T$ and $G = [G_1, G_2, G_3, \dots, G_N]^T$.

Combining the intensities from all the bands linearly, we achieve

$$U = a^T F = [a_1 F_1 + a_2 F_2 + a_3 F_3 + \dots + a_N F_N] \quad (2)$$

$$V = b^T G = [b_1 G_1 + b_2 G_2 + b_3 G_3 + \dots + b_N G_N] \quad (3)$$

where *a* and *b* are the eigenvectors which can be determined using the generalized Eigen system. They can be determined using methods like QR etc. *U* and *V* are the linear combination of vector *F* and *G* and are known as canonical variates. Maximizing the change information requires maximizing variance which requires minimizing the correlation between *U* and *V* under the condition $Var(U) = Var(V) = 1$ with keeping the correlation as positive. Using CCA, searches for linear combinations $U = a^T F$ and $V = b^T G$ of the (ideally) Gaussian distributed variables $[F, G] \in N(\mu, \Sigma)$ with maximum correlation:

$$\begin{aligned} \rho &= Corr(U, V) = \frac{Cov(U, V)}{\sqrt{Var[U][V]}} \\ &= \frac{a^T \sum_{12} b}{\sqrt{a^T \sum_{11} a \quad b^T \sum_{22} b}} \end{aligned} \quad (4)$$

This leads to the coupled generalized Eigen value

problems

$$\begin{aligned} \sum_{fg} \sum_{gg}^{-1} \sum_{gf} a &= \rho^2 \sum_{ff} a, \\ \sum_{fg} \sum_{gg}^{-1} \sum_{gf} a &= \rho^2 \sum_{ff} a \end{aligned} \quad (5)$$

where $\sum_{ff}, \sum_{gf}, \sum_{fg}$ and \sum_{gg} are the covariance matrices between the multidimensional vectors *F* and *F*, *F* and *G*, *G* and *F* and *G* and *G* respectively. MAD Transformation is defined as the subtraction of canonical variates (solutions of Eqns (4) and (5)) in the reverse order.

Two images acquired under different atmospheric conditions with no genuine land cover changes will have a difference component with a nearly Gaussian distribution. Using suitable decision threshold, the no-change pixels can be separated from changed pixels. Furthermore, the sum of squared standardized MAD variates is chi-square distributed, which helps in setting the threshold as

$$\sum_{i=1}^N \left(\frac{MAD_i}{\sigma_i} \right)^2 \ll t \quad (6)$$

As long as the radiometric effects are linear, the pixels chosen can serve as PIFs.

Using PIFs, we are determining radiometric transformation parameters by iteratively reweighted least squares method.

3.2 Change Criteria Function

Principal component analysis takes care of the redundant information in the given bands by incorporating the information from each band by maximizing the information/variance in the first principal component. This could contain both change as well as no-change information. As we aim at detecting significant changes in every band, we use principal component analysis on the difference images (PCD) to maximize the change information only. Change information contained in one component is not replicated in the other ones since the respective directions are orthogonal to each other and the components are uncorrelated. Thus, we have devised a new algorithm to incorporate information from each component such that the most significant change/no-change information can be acquired. For this two criteria are included.

Criteria 1 : If respective band difference images are used for PCA, then redundant no change information can be minimized.

Criteria 2 : If smooth function from mean is defined, then loss of change information is minimum instead of using threshold 1σ or 2σ .

By central limit theorem, it is assumed that the change information is Gaussian distributed. Most of the unchanged or minor changed pixels fall close to the mean of distribution on either side while changed pixels are highly deviated from mean. Usually, the thresholding is strictly compartmentalized with change/no-change binary information, pixels falling outside +/- 3 standard deviation being considered as strictly changed and rest of the others, unchanged. This can lead to valuable loss of change information. In our approach, for every pixel, instead of binary change/no-change information, we devise a fuzzy concept of 'degree of change' which is inversely

proportional to nearness to mean, thus resulting in a smooth inverse triangular change function.

3.3 Contribution of Principal Component Difference

The percentage change information of each component is computed as:

$$Change\ Proportion = \frac{\lambda_i}{(\lambda_1 + \lambda_2 + \dots + \lambda_N)} \quad (7)$$

where λ_i is the Eigen value of the respective PCDs. Starting from the first component, we keep including the lower components in the algorithm till the total percentage variance reaches a given threshold (say, 98 per cent). The rest of the components do not take part in the algorithm as they contain merely 2 per cent information.

3.4 Determination of Inverse Triangular Function

The mean (μ) and standard deviation (σ) of each principal component of difference image is calculated. By virtue of central limit theorem, the addition/subtraction of various random effects would result in the distribution of the PCDs to be nearly Gaussian. Thus, by analyzing each PCD¹⁰ and its corresponding distribution, an inverse triangular function is defined based on the assumption that more distant a pixel value δ from the mean, more likely is that pixel to be categorized as changed. From Figs. 1(a) and 1(b), we can define the inverse triangular function (ITF) as:

$$ITF(\delta)_i = \left\{ \begin{array}{ll} 1 & 0 \leq \delta_i < L_\delta \\ \frac{(\delta_i - \mu_i)}{(L_\delta - \mu_i)} & L_\delta \leq \delta_i < \mu_i \\ \frac{(\delta_i - \mu_i)}{(H_\delta - \mu_i)} & \mu_i \leq \delta_i < H_\delta \\ 1 & H_\delta \leq \delta_i < 255 \end{array} \right\} \quad (8)$$

3.5 Degree of Change

The unified change/no change information (UCM) is acquired using the concept of set-theory as depicted in the

following equations:

$$UCM(\delta)_i = \bigcup_{i=1}^n ITF(\delta)_i \quad (9)$$

which can also be written as

$$UCM(\delta)_i = Max(ITF(\delta)_i), i = 1\ to\ n \quad (10)$$

where $Max(ITF(\delta)_i)$ is the maximum value of inverse triangular function for a given set of PCDs. In Fuzzy System, given N sets of numbers, union of the set at each corresponding point is the maximum value out of all the sets at that point. Since, change information is represented in terms of fuzzy degree of change (DOC)/inverse triangular member function, the union of change information extracted for every pixel from all the available/selected PCDs is the maximum value out of all the values of DOCs for that pixel. Based on the DOC obtained for each pixel from Eqns (9) and (10), classify the pixels as changed if their DOC is greater than a defined value and an unified change mask is generated for all PCD images.

3.6 Change Representation Methodology

Change vector analysis is a technique to analyze changes in terms of magnitude and direction. For this purpose, either features/transformations like greenness, brightness, etc or bands like red, near-infrared etc can be used as input. Changes can be represented according to the magnitude and direction of the change vector. The length of change vector gives the extent of change in the gray level value of pixel whereas the direction gives the information regarding the type of change.

In this paper, two cases have been considered for input to change vector analysis. In *Case (a)*, red and NIR bands have been taken as inputs whereas in *Case (b)*, greenness and brightness as extracted from Kauth-Thomas Transformation have been used. The measurement space with red/ brightness along X-direction and NIR/ greenness along Y-direction is shown in Figs. 2 (a) and 2(b) respectively.

In *Case (a)*, magnitude and direction of change was computed as:

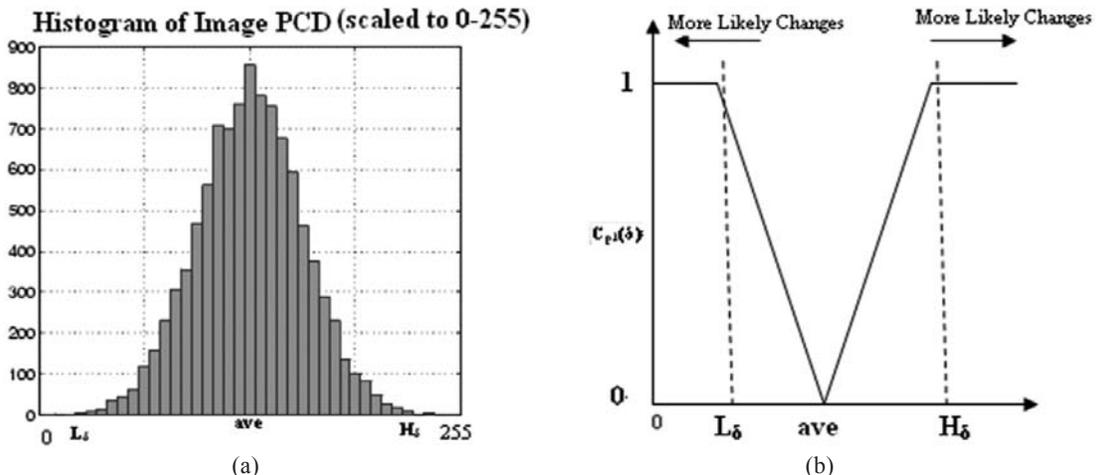


Figure 1. (a) Histogram of the image PCD, and (b) Inverse triangular function of the image PCD.

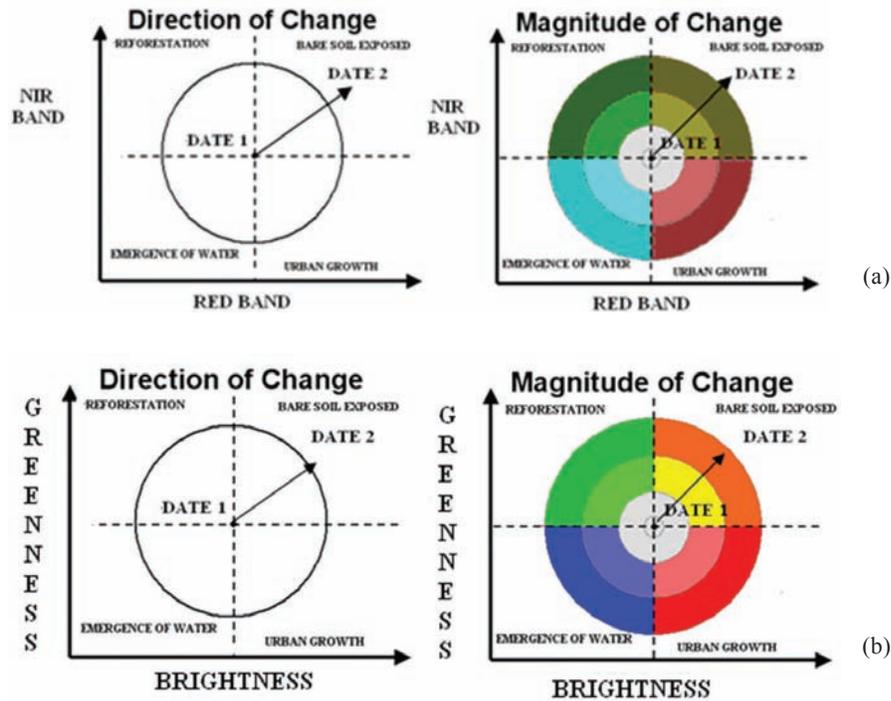


Figure 2. (a) Measurement space with red band along x-axis and NIR band along y-axis and (b) Measurement space with brightness along x-axis and greenness along y-axis.

$$\begin{aligned}
 r &= \sqrt{((\Delta RED_i)^2 + (\Delta NIR_i)^2)} \\
 \theta &= \cos^{-1}\left(\frac{(\Delta RED_i)}{r}\right), \\
 \Phi &= \cos^{-1}\left(\frac{(\Delta NIR_i)}{r}\right)
 \end{aligned} \quad (11)$$

where Δred , ΔNIR are the differences in the spectral values of the corresponding pixels in the two images for the red and NIR bands. r , θ , Φ are the magnitude, direction along red-axis and NIR-axis.

In *Case (b)*, magnitude and direction of change was computed as:

$$\begin{aligned}
 r &= \sqrt{((\Delta Brightness_i)^2 + (\Delta Greeness_i)^2)} \\
 \theta &= \cos^{-1}\left(\frac{(\Delta Brightness_i)}{r}\right), \\
 \Phi &= \cos^{-1}\left(\frac{(\Delta Greeness_i)}{r}\right)
 \end{aligned} \quad (12)$$

where $\Delta Brightness$, $\Delta Greeness$ are the differences in the spectral values of the corresponding pixels in the two images for the greenness and brightness bands for KT-Transformation. The Kauth-Thomas Tasseled Cap transformation rotates the MSS data plane such that the vast majority of data variability is concentrated in two features which are directly related to physical scene characteristics. Brightness is a weighted sum of all the bands, and can be defined in the direction of principal variation in reflectance. It thus measures total reflectance. The

second feature, greenness, is a contrast between the near-infrared bands and the visible bands. High Greenness values represent targets with high densities of green vegetation, while the flatter reflectance curves of soils are expressed in low Greenness values. Further, it is a linear transformation maintaining the affine relationship between the raw data. The transformation is invariant and consistent over different data sets¹¹. Thus, r , θ , Φ are the magnitude, direction along greenness-axis and brightness-axis. Since we have use IRS-ID data here, the Kauth-Thomas greenness- brightness coefficients¹² becomes Eqns (13-14),

$$\begin{aligned}
 \text{Greenness} &= -0.4233 * \text{Green} - 0.4328 * \text{Red} \\
 &+ 0.6490 * \text{Nir} - 0.4607 * \text{Swir} \quad (13)
 \end{aligned}$$

$$\begin{aligned}
 \text{Brightness} &= 0.1072 * \text{Green} + 0.1219 * \text{Red} \\
 &+ 0.4341 * \text{Nir} + 0.8861 * \text{Swir} \quad (14)
 \end{aligned}$$

For the each pixel classified as changed by unified change mask, the change vector magnitude and direction was determined for both the cases. These changed pixels are represented by pseudo coloring scheme shown in Figs. 3(a) and 3(b) for *Case (a)*. The same procedure was adopted for *Case (b)* and shown in Figs. 4(a) and 4(b). The pseudo coloring scheme change direction for different categories of changes while magnitude for sub categories

4. RESULTS AND DISCUSSION

Figures 3(a) to 3(d) show MSS Image of date 1, MSS image of date 2, change/no-change map using the proposed thresholding and change mask using *Case (a)* CVA (red v/s NIR band) with the proposed thresholding respectively.

Figures 4(a) to 4(d) show MSS Image of date 1, MSS image of date 2, change/no-change map using the proposed

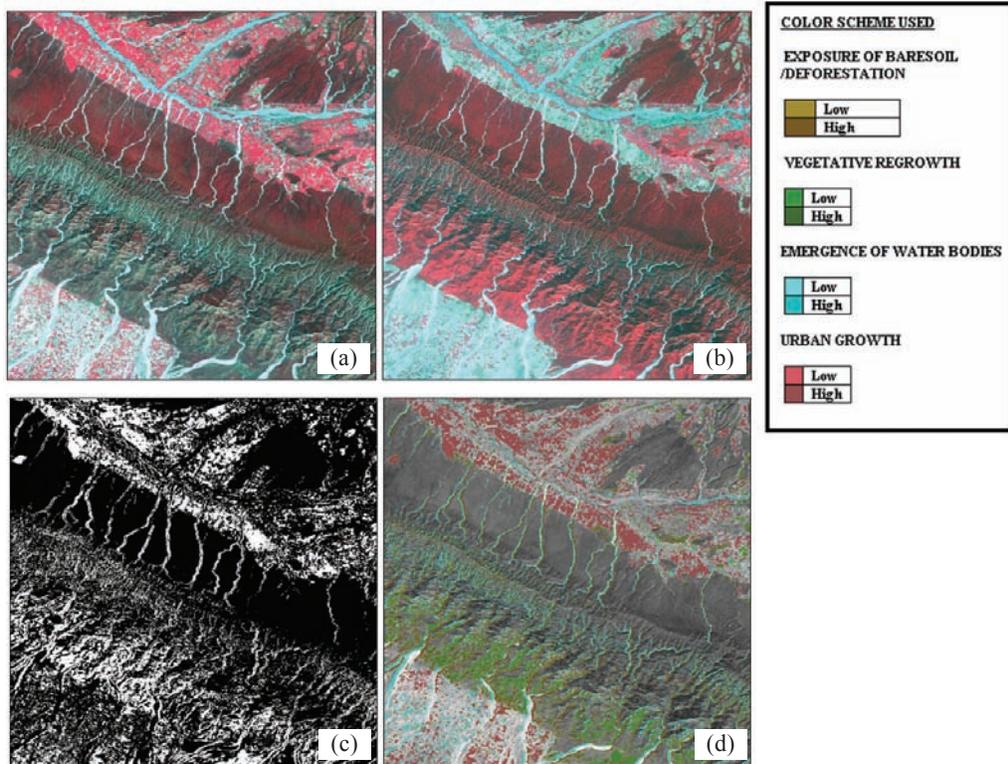


Figure 3. (a) Image of date 1, (b) image of date 2, (c) change mask generated by enhanced thresholding, and (d) change mask using change vector analysis case (a).

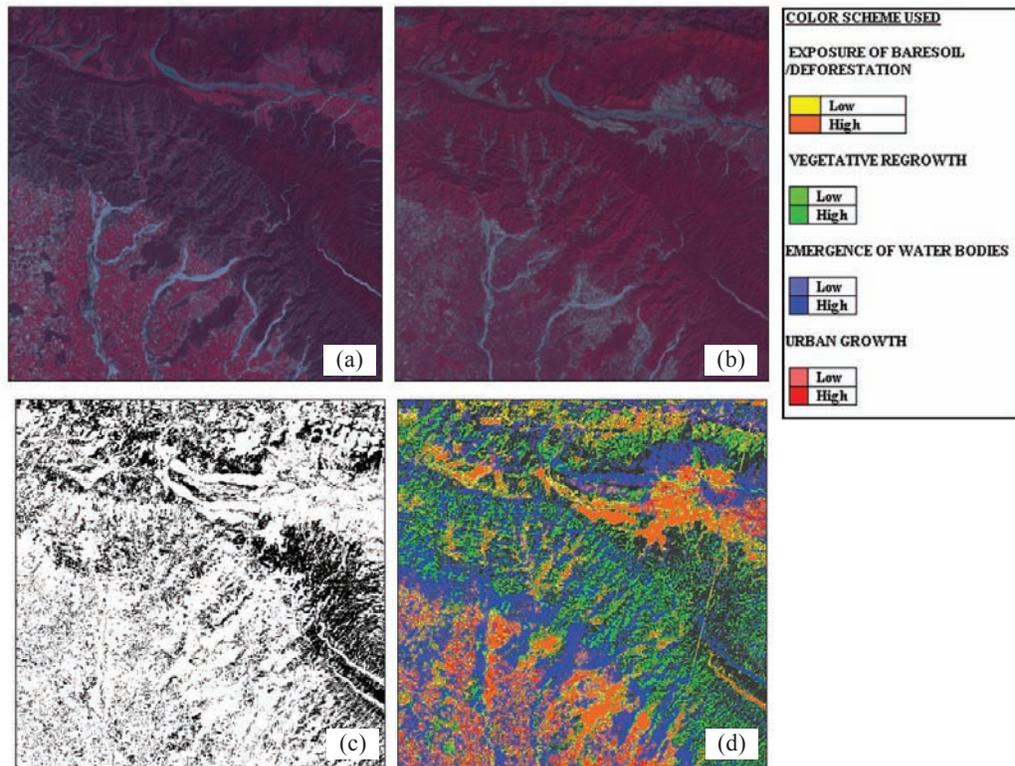


Figure 4. (a) Image of date 1, (b) image of date 2, (c) change mask generated by enhanced thresholding, and (d) change mask using change vector analysis case (b).

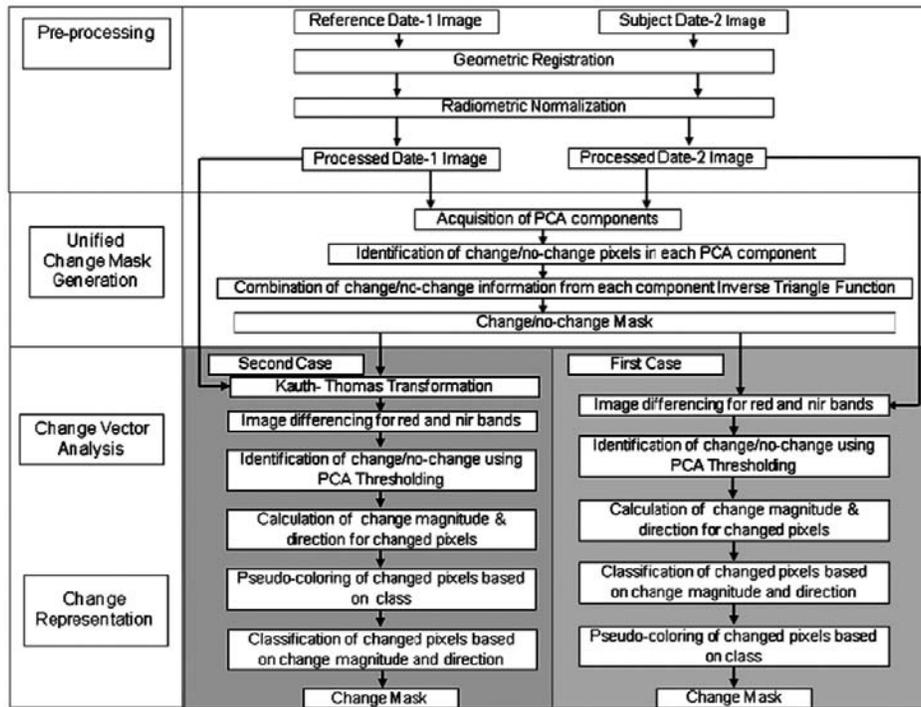


Figure 5. Block diagram depicting the improved CVA technique using enhanced PCA and inverse triangular function-based thresholding.

thresholding and change mask using *Case (b)* CVA (brightness v/s greenness) with the proposed thresholding respectively.

Accuracy of the method was performed using quantitative error assessment methods like error matrix and KAPPA analysis. By means of reference data, the images were classified using supervised classification and class based changes were compared with the results from the proposed method. The error matrix is shown in Table 1. KAPPA coefficient was found to be 0.5271 and the accuracy from the error matrix came out to be 74.89 per cent.

5. SUMMARY

Change Vector Analysis can intuitively detect the nature of changes which have taken place in multi-temporal multispectral satellite imagery but it suffers poorly from improper thresholding. By using principal component analysis on bands, high variance component may contain both change and no-change, thus the capabilities of PCA in maximizing change information are not utilized to the best extent. The proposed technique has taken care of these issues. By incorporating principal component analysis and inverse triangular functions, thresholding technique can be improved. Further, PCA has been used here on difference bands, which helps to maximize only change information in high variance components. Also, instead of using a strictly compartmentalized thresholding with change/no-change binary information, a fuzzy concept of ‘degree of change’ has been introduced, which is inversely proportional to nearness to mean, thus resulting in a smooth inverse triangular change function. Thus, according to the user-defined thresholding, pixels with a given degree of change can be chosen. Also, a union of all the selected components of PCD (according to the percentage variance) makes sure that

Table 1. Error matrix for the proposed method

Proposed Method	Reference Data	
	No Change	Change
No Change	237564	71256
Change	167868	475762

change information from all the components can be integrated efficiently. This, combined with change vector analysis, can integrate the change information from all the bands as well as classify the nature and magnitude of changes.

ACKNOWLEDGEMENT

The authors wish to express their sincere gratitude to the Director, Defence Electronics Applications Laboratory, Dehra Dun for providing the necessary support to carry out the research and to publish the work.

REFERENCES

1. Kolehmainen, Karoliina & Ban Yifang. Multitemporal spot images for urban land-cover change detection over Stockholm between 1986 and 2004. *Int. Archives Photogrammetry, Remote Sensing Spatial Info. Sci.*, 2008, **37**(B6), 63-66.
2. Kuzera, Kristopher; John Rogan & Eastman, J.R. Monitoring vegetation regeneration and deforestation using change vector analysis: Mt. St. Helens study area. *In the Annual Conference Baltimore, ASPRS 2005, Maryland, 7-11 March 2005.*
3. Chen, J.; Gong, P.; He, C.; Pu, R. & Shi, P. Land use/ land cover change detection using change vector analysis.

- Photogrammetric Engg. Remote Sensing*. 2003, **69**(4), 369-79.
4. Sohl, T.L. Change Analysis in United Arab Emirates: An investigation of Techniques. *Photogrammetric Engg. Remote Sensing*. 1999, **65**(4), 475-8
 5. Smith, L., A tutorial on principal components analysis. <http://csnet.otago.ac.nz/cosc453/studenttutorials/principal> [Accessed on 30 June 2010].
 6. Shlens, J. A Tutorial on principal components analysis. <http://www.dgp.toronto.edu/~aranjan/tuts/pca.pdf> [Accessed on 30 June 2010].
 7. Baisantry, M.; Negi, D.S. & Manocha, O.P. Automatic relative radiometric normalization for change detection of satellite imagery. In the Proceedings of International Conferences on Advances in Computer Science, New Delhi, 20-21 December 2011. pp. 42-45.
 8. Canty, M.J.; Nielsen, A.A. & Schmidt, M. Automatic radiometric normalization of multi-temporal satellite imagery. *Remote Sensing Environment*, 2004, **91**, 441-51.
 9. Hardoon, D.R.; Szedmak, S. & Taylor, J.S. Canonical correlation analysis; An overview with application to learning methods. Department of Computer Science, Royal Holloway, University of London. Technical Report: CSD-TR-03-02, 2003.
 10. Gong, P. Change detection using principal component analysis and fuzzy set theory. *Canadian J. Remote Sensing*. 1993, **19**(1), 22-29.
 11. Crist, E.P. & Cicone, R.C. A Physically-based transformation of thematic mapper data-The TM tasseled cap. *IEEE Tran. Geosc. Remote Sensing*, 1984, **22**(3), 256-63.

12. Nag, P. & Kudrat, M. Digital remote sensing. Concept Publishing Company, New Delhi, 1998. pp. 160-62.

Contributors



Ms Munmun Baisantry obtained her BE (Electronics and Communications) from Netaji Subhas Institute of Technology in 2009. Presently, working as a Scientist in the Defence Electronics Applications Laboratory (DEAL), Dehra Dun. Her research interests include : Radiometric normalization, image change detection, image fusion and object-based image analysis.



Dr Devendra Singh Negi received his MSc (Mathematics) from DAV Post Graduate College, Dehra Dun in 1987. Received his PhD in Applied Mathematics (Image Processing) from Garhwal University, Srinagar, Uttaranchal in 2007. His research interests include : Object-based image analysis, applications development. He is a Life member of Indian Society of

Remote Sensing.



Mr Om Prakash Manocha received his MSc (Physics) from DAV Post Graduate College, Dehra Dun in 1982. His research interests include : Object-based image analysis, change detection, and analysis application and image fusion. He is Life member of Indian Society of Remote Sensing, Life Fellow of IETE.