A Systematic Study of Change Detection Algorithms

K.Venkateswaran
Assistant Professor
Department of ECE
Kongu Engineering College
Erode

N.Kasthuri, Ph.D
Professor
Department of ECE
Kongu Engineering College,
Erode

D.Dorin Jeni
PG Scholar
Department of ECE
Kongu Engineering College
Erode

ABSTRACT
The great presumption of change detection has led to rapid development of diverse change detection algorithms. Unsupervised change detection has a vital role in a wide variety of applications like remote sensing, motion detection, environmental monitoring, medical diagnosis, damage assessment, agricultural surveys, surveillance etc. In this paper a systematic survey of the commonly used methodologies for unsupervised change detection is presented.

General Terms
Remote Sensing, Image Processing, Pattern Recognition.

Keywords
Change detection, image differencing, feature vector, thresholding, clustering.

1. INTRODUCTION
The exploitation of multitemporal images has been prudently done in change detection. In the change detection process two images of the same area are taken at two different time instances and then these multitemporal images are processed to identify the changes that may have taken place between the time instances in which the two images were taken. The output of the change detection process is a binary change map, indicating the location of the changes.
Change detection methods can be extensively classified into supervised and unsupervised paradigms based on the nature of method used for processing the images. In supervised change detection, a training set is required for the learning process of the classifiers. This training set is obtained from the ground truth data [1].The latter technique performs change detection by using only the multitemporal images and in this method no ground truth is needed [2].

The process of change detection is of widespread interest since it is having a wide variety of applications in diverse domains like remote sensing[3], motion detection[4], video surveillance[5], damage assessment[6], agricultural surveys[7], environmental monitoring[8], analysis of urban changes[9], medical diagnosis[10].

A wide variety of change detection techniques have been introduced in the literature. Due to the massive growth of the geographic database, it is more practical to focus on the unsupervised approach than the supervised one. Currently, many unsupervised change detection techniques have been proposed. Some of these are image algebra, transformation of the multitemporal images, image classification, advanced models, Geographical Information System (GIS) approaches etc. In image algebra, a certain algebraic technique is adopted to generate the difference image from the multitemporal input images [10], and then the difference image thus obtained is processed in order to identify the changes. The problem of differentiating the changed pixels from unchanged pixels can be regarded as an image segmentation problem. Generally, the most popular solution to this problem is thresholding.

In section Ⅱ the problem of unsupervised change detection is defined. Section Ⅲ provides description of some of the preprocessing steps. In section Ⅳ some techniques used to generate difference image is discussed. In section V and VI thresholding and clustering techniques used for generating the change map is detailed. Finally in section VII conclusion is drawn.

2. PROBLEM STATEMENT
In order to perform change detection stringently, let $\{X_1, X_2\}$, be two images of the same area taken at two different time instances. Here every pixel, P in each image, $P \in R^l$ and the intensity$I(P) \in R^k$. Depending on the value of k images can be either gray scale(k =1),RGB colour images(=k 3),multispectral images(k >10),hyperspectral images(k >100). In change detection algorithm, an image sequence is taken as input and the output generated is a binary image $B: R^l \rightarrow [0,1]$, called a change map.

3. PREPROCESSING TECHNIQUES
The main aim of unsupervised change detection is to emphasize the important changes and to deemphasize the unimportant changes. The following section elucidates the preprocessing steps for deemphasizing the unimportant changes. Radiometric adjustments and image registration are indispensable steps in change detection process.

3.1 Radiometric adjustments
Radiometric adjustment is done in order to precompensate the inconsistency between intensity variations caused by strength or position of light sources, values surveyed by spectral reflectivity, different imaging seasons, different cover areas of cloud, rain or snow etc. Hence it is necessary to perform radiometric corrections before change detection [11]. To rectify the radiation distortion which is irrelevant to the radiating features at the surface, state of sensors, dispersion, solar illumination, absorption of atmosphere etc., absolute radiometric correction is employed. The commonly used technique involves modifying the radiation value to the standard reference value with spectral curves but this technique is costly and is not practical to set the reference values based on spectral reflectance curve hence it is tedious to implement the absolute radiometric correction in practice.
Illumination invariant change detection utilizes intensity normalization technique [12]. In image normalization the

pixel intensity values in one image are normalized so that the mean and variance of both the images remains same. $\widetilde{X_2}(x) = \frac{\sigma_1}{\sigma_2}\{X_2(x) - \mu_2\} + \mu_1$where, $\widetilde{X_2}(x)$is the normalized value of second image and σ_i and μ_i are the standard deviation and mean values of intensity. When both the input images are normalized inorder to obtain zero mean and unit variance then the decision thresholds become independent of the intensity values. Normalization can also be done using block processing technique.

The suppression of irrelevant changes due to shadows of objects can be done for surveillance application [13] using $\left({G}/{(R+G+B)}\right), \left({B}/{(R+G+B)}\right), (R+G+B)$these values of coordinates are more appropriate than the values of RGB coordinates. Radiometric pre-processing can also be done by knowing the accurate location of sun and to fit models to images this removes the shadows of buildings as an effect of direct light from the sun before applying change detection algorithm [14]. False changes due to speckle noise can be suppressed using frame averaging [15-17].

3.2 Image Registration

The multi-temporal images must be co-registered before applying change detection algorithm in order to avoid spurious results of change detection. Co-registration is the process of arranging both the input images into the same coordinate frame. If misregistration occurs large number of false changes occurs due to displacement of image.

Registration is done by using low dimensional spatial transforms like similarity and projective transforms [18-21] software implementations of such algorithms are also available. Many recent algorithms for registration are having the ability to switch automatically to higher order transformations after being initialized with a low order similarity transformation [22]. A nonglobal transformation is needed to determine the corresponding changes between images when the objects in the scene are deformable [23]. Localized registration errors may occur in change detection process even though the global error measures are considered as the best, hence the designer should take in to account the possibility of occurrence of localized errors. The effects of registration errors are analysed by D.A.Stow et.al [24-26]. Hence, for obtaining an error less than 10% one-fifth pixel registration accuracy is needed. Please use a 9-point Times Roman font, or other Roman font with serifs, as close as possible in appearance to Times Roman in which these guidelines have been set. The goal is to have a 9-point text, as you see here. Please use sans-serif or non-proportional fonts only for special purposes, such as distinguishing source code text. If Times Roman is not available, try the font named Computer Modern Roman. On a Macintosh, use the font named Times. Right margins should be justified, not ragged.

4. IMAGE DIFFERENCING

In the literature, various supervised and unsupervised methods for the detection of changes in remote sensing images have been proposed. In this paper, a survey of some of the commonly used unsupervised change detection algorithms has been done. From a methodological viewpoint, unsupervised change detection involves two key steps. The first key step is to produce a difference image from the multitemporal images to enunciate the changed regions. In the second step, the difference image is analyzed in order to distinguish between the changed and unchanged regions. This can be considered as an image segmentation problem.

Among the earliest methods of getting the difference image is the commonly used one is simple differencing. The simple difference image is obtained by,

$D(X) = D(X_2) - D(X_1)$.

Another technique used to obtain difference image is Change Vector Analysis (CVA) [27] where, for all the spectral channels for every pixel in the image, a feature vector is generated. In CVA, for generating the difference image, the modulus of the difference between the feature vectors of the multitemporal images is taken. This technique is commonly used in multispectral images.

Image differencing can also be generated by image ratioing. Ratioing is considered to b a precise means of determining the areas of change. In image ratioing, the ratio of the intensities in the multitemporal images is taken. The difference image

$$D(X) = \frac{I(X_2)}{I(X_1)}$$

In areas of change, the value of ratio could be either greater than one or less than one depending upon the nature change [28]. Since the image ratioing is based on non-normal distribution its use in change detection has been criticized.

In some recent algorithms for change detection, before performing image segmentation (thresholding or clustering) we perform feature extraction on the difference image by utilizing principal component analysis(PCA), discrete wavelet transform (UDWT), stationary wavelet transforms (SWT), discrete wavelet transform (DWT)etc.,

In [29] the change map is obtained by combining principal component analysis (PCA). In this transformation the first principal component is having the largest possible variance , which accounts for the major percentage of the variability in the data, and each succeeding component in turn has the highest variance possible under the condition that it must be orthogonal to the preceding components, so that they are uncorrelated. PCA is applied to extract the eigen vectors from the non overlapping blocks of the difference image, then an $h \times h$ neighbourhood data is projected into the eigen vector space for extracting the feature vector corresponding to each pixel in the difference image. The average vector of the $h \times h$ neighbourhood data in the difference image is defined by

$$\psi = \frac{1}{M}\sum_{p=1}^{M} x_d^p$$

where x_d^p represents the difference vector $x_d(y, x)$, while p represents an index with $1 \leq p \leq M = \lfloor (H \times W)/h \times h \rfloor$.

The covariance matrix is defined by

$$C = \frac{1}{M}\sum_{p=1}^{M} \Delta_p \Delta_p^T$$

where $\Delta_p = x_d^p - \psi$. The feature vector space for each pixel at spatial location (i, j) is given by

$$v(i,j) = [v_1\ v_2\ ...\ v_S]^T$$

where $1 \leq S \leq h^2$. Now the feature vector space can be used to detect the changes by performing clustering.

Since multi scale data fusion is not considered in the PCA based approaches, it is susceptible to false alarm. Hence in [30] PCA has been replaced with UDWT. Here undecimated discrete wavelet transform is used to perform multiscale decomposition of the difference image. The intra scale feature vector at spatial location (i,j) is obtained as,

$v_k(i,j) = [\Delta_k(n_1(i,j)) \ldots \Delta_k(n_N(i,j))]$ where $\Delta_k = \{\Delta_1\, \Delta_2 \ldots \Delta_{3S+1}\}$, $n(i,j) = \{n_1(i,j), n_2(i,j), \ldots n_N(i,j)\}$,and N is the total number of neighbouring pixels.

The performance of multiscale decomposition of the difference image can be improved by replacing UDWT with SWT. In [31] mean shift algorithm is performed on the difference image to minimize the noise. The feature vectors are extracted from the difference image by using SWT. The new feature vector is extracted at each spatial location (i,j) as

$$v(i,j) = \left\{ \begin{matrix} M'(i,j), M'_{1,A}(i,j), M'_{1,A}(i,j), M'_{1,H}(i,j), M'_{1,V}(i,j), M'_{2,A}(i,j), \\ M'_{2,H}(i,j), M'_{2,V}(i,j) \end{matrix} \right\}$$

Another method for generating a difference image is by integrating the merits of different algorithms used to generate the difference image. In this algorithm fusion techniques are used for merging multiple difference images. Hence a more consistent difference image can be obtained [32].Also, in [33], image fusion is used to gather information collectively from a log-ratio and mean-ratio image. Multiscale transforms such as discrete wavelet transform (DWT), has been used widely for image fusion. By using Discrete Wavelet Transform detailed information can be easily extracted from the images, since it provides localization of frequencies in both time and space domains. DWT can be replaced with other multiscale transforms like contourlets , curvelets etc, as these are having better shift-invariance and directionality than wavelets.

5. THRESHOLDING

Once a difference image has been obtained, then the problem of change detection reduces to an image segmentation problem. The most popular solution for image segmentation is thresholding. The output of the thresholding process done on the difference image gives a binary change map, in which one state indicates the changed pixels and the other state represents the unchanged pixels. Thresholding can be classified on the basis of parameters like: shape of the histogram, entropy, object-attribute, clustering mechanism etc. Some of the popular and efficient thresholding techniques are discussed in this paper.

5.1 Thresholding based on shape of the histogram

In convex hull thresholding, [34], concavities of the histogram $h(g)$ represented by the convex hull, $Hull(g)$ given by, $h(g) = |Hull(g) - p(g)|$, where $p(g)$ is the input image distribution. In this method, after determining the convex hull, the point having the deepest concavity is selected as the thresholding level. If there is more than one such point, the point having low busyness of the threshold image edges can be taken as the thresholding level.

In peak and valley thresholding [35], the peak analysis of the histogram is done. Peak analysis is done by the convolving the with a differencing and smoothing kernel. The smoothing aperture of the kernel is adjusted to get a histogram having two-lobe function. By performing the differencing operation on the kernel, a triplet of incipient, maximum, and terminating zero crossings of the histogram lobe $S = [(e_i, m_i, s_i), i = 1, ..2]$.And hence the thresholding level should be between the first terminating and second initiating zero-crossing that is:

$$T_{opt} = \gamma e_1 + (1-\gamma)s_2, 0 \leq \gamma \leq 1$$

In shape-modelling thresholding [36], a simple functional approximation to the PMF consisting of a two step function is used. Thus we obtain the sum of squares between a bi-level function and the histogram is minimized. So we can obtain the thresholding solution, T_{opt}, by using iterative search.

5.2 Minimum error thresholding

Another commonly used thresholding approach is minimum-error thresholding. In minimum error thresholding, the image is considered to be a mixture of background and foreground pixels:$p(g) = P(T).p_f(g) + [1 - P(T)].p_b(g)$. This can be performed in many ways. For example, Lloyd, [37] considered the equal variance Gaussian density function which performs minimization of the total misclassification error iteratively. While Kittler and Illingworth [38] used a minimum error Gaussian density-fitting functions. More recently Cho et.al [39], proposed an enhancement of this thresholding technique, by concluding that the means and variances determined from the truncated distributions result in a bias. But this bias is relevant only when the two histogram modes are not distinguishable.

5.3 Entropy based thresholding

This kind of thresholding algorithms are based on the entropy of the distribution of the gray levels in a scene. Here, maximum information transfer is considered to take place by maximizing the entropy of the thresholded image [40]. Other methods perform minimization of the cross-entropy between the input gray level image and the output binary image, as a measure of preservation of information [41] or a measure of fuzzy entropy [42].

In entropic thresholding [43], the foreground and background pixels of the image are considered as two different signal sources, and hence when the sum of the two class entropy reach its maximum, the thresholding is set to be optimum. In [44], the entropic correlation is obtained as:

$$TC(T) = C_b(T) + C_f(T)$$
$$= -log\left[\sum_{g=0}^{T} \frac{p(g)}{P(T)}\right]^2 \left[\sum_{T+1}^{G} \frac{p(g)}{1-P(T)}\right]^2$$

And obtain the threshold that maximizes it.

5.4 Thresholding based on attribute similarity

In this kind of thresholding algorithms, the thresholding value is selected based on some attribute quality or similarity measure between the original image and the binarized version of the image.

The attributes taken can be either edge matching [45],shape compactness [46], gray level moments[47],connectivity [48] or texture[49].In some other algorithms, the similarity between the original gray-level image and binary image using fuzzy measure [50],or resemblance of the cumulative probability distributions [51].

5.5 Locally adaptive thresholding

In this class of thresholding, the thresholding value is calculated at each pixel depending on some local statistics like range, variance, or surface-fitting parameters of the pixel neighbourhood. Thus the threshold $T(i,j)$ is indicated as a function of the coordinates (i,j) of each pixel [52], or else the determination of foreground or background are indicated by a logical variable $B(i,j)$. The parameter can be the local variance [53], local contrast [54], or center-surround scheme [55].

6. CLUSTERING MODELS

The most common image thresholding solution to the image segmentation problem is image thresholding. But thresholding has many drawbacks. The success of the thresholding algorithm depends on the statistical distribution of the changed and unchanged classes, but it is difficult to calculate this statistical distribution a single threshold level accurately. Also the noise cannot be precisely calculated by using a single threshold level. In [56] two techniques based on the Bayes theory are proposed. In one technique the decision threshold is automatically selected by using Expectation Maximization (EM) which maximizes the change detection error assuming that the pixels of the difference image are statistically independent. In the second approach which is based on Markov Random Field (MRF) the difference image is analysed by determining the spatial information which is available in the neighbourhood of each pixel. In [57] a 2D-OTSU segmentation which is less sensitive to noise is introduced. This is done by combining the spatial information of the neighbourhood.

6.1 Clustering by thresholding

Otsu [58] proposed suggested an optimum threshold for detecting the changes by minimizing the weighted sum of within class variance of the foreground and background pixels. This method gives better results when the numbers of pixels in both classes are equal. This technique is still one of the most commonly used techniques for thresholding.

6.2 K-means clustering

k-means is one of the most commonly used algorithm for unsupervised learning for solving clustering problem. The main aim is to determine the centroids for each cluster. The centroids must be placed in an appropriate location since inappropriate locations cause wrong results. The next step is to assign each point belonging to a given data set to its nearest centroid. When all the points in the given data set have been associated to a nearest centroid the first step is completed and an early classification is done. Then the new k centroids of the clusters are recalculated and a new merging has to be done between the same data set and the nearest new centroids. A loop has been generated. Thus k centroids change their location step by step until no more changes are done. The k-means algorithm aims at minimizing the objective function,

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i^{(j)} - c_j \right\|^2$$

where $\left\| x_i^{(j)} - c_j \right\|^2$ is the distance measured between the data point $x_i^{(j)}$ and the cluster centre c_j, n is the number of data points.

In [29] the change map is obtained by combining principal component analysis (PCA) and k-means clustering (k=2). First PCA is applied to extract the eigen vectors from the non overlapping blocks of the difference image, then an $h \times h$ neighbourhood data is projected into the eigen vector space to extract the feature vector for each pixel in the difference image. Now the feature vector space is clustered into two by using k-means algorithm. The mean feature vector represents each cluster. Then for every pixel the minimum Euclidean distance between its feature vector and the mean feature vector of the clusters is determined to get the change map.

In [59] the difference image is decomposed into different side bands using wavelet transform to get the feature vectors for each pixel and the feature vectors are clustered by using k-means algorithm.

In [60] in order to consider the non-linearities in the input images the k-means algorithm is modified using a kernel function. The kernel based procedure is interpreted as mapping the data from original input space into a potentially higher dimensional feature space where linear method may then be used to distinguish changed from unchanged pixels. The feature space spanned by kernel function maximizes separability between the changed and unchanged pixels. The formation of difference image in feature space is done using kernel mapping function.

6.3 Fuzzy C-means clustering

One of the most popularly used soft clustering technique is fuzzy c-means clustering in which each element in a data set belongs to two or more clusters. It aims at the minimization of the objective function,

$$J = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^{m} \left\| x_i - c_j \right\|^2, 1 \leq m < \infty$$

where m is any real number greater than 1,u_{ij} is the degree of membership of x_i in the cluster j. x_i is the $i^{th}$of d-dimensional measured data. c_j is the d-dimension center of the cluster. $\|*\|$ is any norm expressing the similarity between any measured data and the center.

Among the soft clustering techniques, FCM [61] is the most commonly used one because of its robustness and detail-preserving characteristics than other hard clustering techniques [62].Although the conventional FCM works well on almost all images, since FCM does not take into account the spatial information ,it is very sensitive to noise. Hence for change detection in synthetic aperture radar images, conventional FCM does not provide good results. In SAR images a modification of FCM which takes into account the spatial information named fuzzy local information c-means clustering is used [33].

7. ACCURACY ASSESMENT AND PERFORMANCE EVALUATION

The evaluation of performance of change detection algorithms can be done visually and quantitatively. The quantitative analysis can be done using the ground truth reference and check rules. The general approach to obtain ground truth reference is to perform field survey with the assistance of historical GIS data.

The visual analysis can be done using flicker animation [63] where the registered images are played rapidly with an interval of about a second each. When changes occur the changed regions seem to flicker else one perceives a steady image. In this technique, the change mask is estimated based on the effect of flickering.

To assess the accuracy quantitatively in the absence of ground truth reference consistency check rules based on rationality evaluation with post classification comparison [64] is also used.

There are several standard techniques by which the binary change map can be combined to the ground truth provided the ground truth has been established. It can be done by using the following quantities.

The binary classifier generates four outcomes. If the predicted output and the actual value is *p*, then it is considered a *true positive* (TP); but if the predicted output is *p* actual value is *n* then it is regarded to be a *false positive* (FP). On the other hand, a *true negative* (TN) occurs when both the predicted output and the actual value are *n*, and *false negative* (FN) occurs when the predicted output is *n* while the actual value is *p*.

Rosin [65] described the following methods by which a classifiers performance can be quantified:

1) percentage correct classification

$$PCC = (TP + TN)/(TP + FP + TN + FN)$$

2) Jaccard coefficient

$$JC = TP/(TP + FP + FN)$$

3) Yule coefficient

$$YC = |TP/(TP + FP) + TN/(TN + FN) - 1|$$

4) accuracy

$$ACC = (TP + TN)/(p + n)$$

For a classifier with tunable parameters the receiver operating characteristics is investigated. A ROC space is defined by FPR and TPR as *x* and *y* axes respectively, which depicts relative trade-offs between true positive (benefits) and false positive (costs). Since TPR is equivalent with sensitivity and FPR is equal to (1−specificity), the ROC graph is sometimes called the sensitivity vs. (1 − specificity) plot.

In order to assess the accuracy, kappa statistic is used. Kappa statistic is a measure of accuracy based on agreement between, error matrix and chance agreement [66]. If the change detection map and the reference image are in complete agreement, then the kappa value is 1. If there is no agreement among the change detection map and the reference image, the kappa value is 0.

Since the input images are subjected to various noises during acquisition, it is necessary to evaluate the robustness of the change detection algorithm to noise[67]. It can be done by plotting the values of PSNR and tau which can be determined as follows:

$$PSNR = 10log_{10}(\frac{255}{\sum_{i=1}^{M}\sum_{j=1}^{N}(x(i,j) - \hat{x}(i,j))^2})$$

Where M and N are number of rows and number of columns of the input image, $x(i,j)$ is the input image in the absence of noise and $\hat{x}(i,j)$ is the noisy version of the input image.

$$\tau = 1 - \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}|C_1(i,j) - C_2(i,j)|}{MN}$$

Where $C_1(i,j)$ is the change map generated from the input images in the absence of noise and $C_2(i,j)$ is change map generated by using the noisy version of the input images.

8. CONCLUSION

Change detection finds application in diverse disciplines like remote sensing video surveillance etc., In this paper a survey of the techniques which have been used and also some of the recently used techniques is discussed in this paper. We hope that our survey of change detection algorithms , provides an insight to an algorithm designer.

9. REFERENCES

[1] S. M. Metev and V. P. Veiko, *Laser Assisted Microtechnology*, 2nd ed., R. M. Osgood, Jr., Ed. Berlin, Germany: Springer-Verlag, 1998.

[2] L.Bruzzone and D.F.Prieto, "Automatic analysis of the difference image for unsupervised change detection," *IEETrans.Geosci.RemoteSens*.,vol.38,no.3,pp.1171-1182,2000

[3] L.Bruzzone and D.F.Prieto, "An adaptive semiparametric and context-based approach to unsupervised change detection in multitemporal remote-sensing images,"*IEEE Trans.*Image Processing.,vol.11,no.4,pp.66-77,1996.

[4] C. Dumontier, F. Luthon, and J.-P. Charras, "Real-time DSP implementation for MRF-based video motion detection," IEEE Trans. Image Process., vol. 8, no. 10, pp. 1341–1347, Oct. 1999.

[5] R. Collins,A. Lipton, and T.Kanade, "Introduction to special section on video surveillance," *IEEE Trans.Pattern Anal.Mach.Intell*.,vol 22,no.8,pp.745-746,Aug.2000.

[6] K.Grover,S.Quegan and C. da Costa Freitas, "Quantitative estimation of tropical forest cover by SAR," IEEE Trans. Geosci. Remote Sens., vol.37, no. 1, pp. 479–490, Jan. 1999.

[7] L.Bruzzone and S.B.Serpico, "An iterative technique for the detection of land-cover transitions in multitemporal remote- sensing images," *IEEE Trans.Geosci.Remote Sens*.,vol.35,,pp.858-867,July 1997.

[8] P.S Chavez,Jr and D.J MacKinnon, "Automatic detection of vegetation changes in the south western united states using remotely sensed images,"*Photogramm.Eng.Remote Sensing,vol.60,no.5,pp.1285-1294,1994.*

[9] K. R. Merril and L. Jiajun, "A comparison of four algorithms for change detection in an urban environment,*" Remote Sens. Environ., vol. 63, no. 2, pp. 95–100, Feb. 1998.*

[10] T. Celik, "A Bayesian approach to unsupervised multiscale change detection in synthetic aperture radar images," *Signal Process.*, vol. 90, no. 5,pp. 1471–1485, May 2010.

[11] Leonardo.P,Franceisco.G,Jose.A,SobrinoJuan.C,Jimenez .M and Haydee.K,2006,"Radiometric correction effects in lansat muti-date/multi-sensor change detection studies,International Journal of Remote Sensing,2,pp.685-704.

[12] X. Dai and S. Khorram, "The effects of image misregistration on the accuracy of remotely sensed change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 5, pp. 1566–1577, Sep. 1998.

[13] A. Elgammal, R. Duraiswami, D. Harwood, and L. S. Davis, "Background and foreground modeling using nonparametric kernel density estimation for visual surveillance," *Proc. IEEE*, vol. 90, no. 7, pp.1151–1163, Jul. 2002.

[14] S. Watanabe, K. Miyajima, and N. Mukawa, "Detecting changes of buildings from aerial images using shadow and shading model," in *Proc. ICPR*, 1998, pp. 1408–1412.

[15] L. M. T. Carvalho, L. M. G. Fonseca, F. Murtagh, and J. G. P.W. Clevers, "Digital change detection with the aid of multi-resolution wavelet analysis,"*Int. J. Remote Sens.*, vol. 22, no. 18, pp. 3871–3876, 2001.

[16] S. Fukuda and H. Hirosawa, "Suppression of speckle in synthetic aperture radar images using wavelet," *Int. J. Remote Sens.*, vol. 19, no. 3, pp. 507–519, 1998.

[17] S. Quegan and J. Schou, "The principles of polarimetric filtering," *in Proc. IGARSS, Aug*. 1997, pp. 1041–1043.

[18] L. G. Brown, "A survey of image registration techniques," *ACM Comput. Surv.*, vol. 24, no. 4, 1992.

[19] S. Lavallee, Registration for Computer-Integrated Surgery: Methodology, *State of the Art. Cambridge, MA: MIT Press*, 1995.

[20] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Med. Image Anal.*, vol. 2, no. 1, pp. 1–36, 1998.

[21] B. Zitová and J. Flusser, "Image registration methods: A survey," *Image Vis. Comput.*, vol. 21, pp. 977–1000, 2003.

[22] C. V. Stewart, C.-L. Tsai, and B. Roysam, "The dual-bootstrap iterative closest point algorithm with application to retinal image registration," *IEEE Trans. Med. Imag.*, vol. 22, no. 11, pp. 1379–1394, Nov. 2003.

[23] J. Barron, D. Fleet, and S. Beauchemin, "Performance of optical flow techniques," *Int. J. Comput. Vis.*, vol. 12, no. 1, pp. 43–77, 1994.

[24] D. A. Stow, "Reducing the effects of misregistration on pixel-level change detection," Int. J. Remote Sens., vol. 20, no. 12, pp. 2477–2483, 1999.

[25] J. Townshend, C. Justice, C. Gurney, and J. McManus, "The impact of misregistration on change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 5, pp. 1054–1060, Sep. 1992.

[26] X. Dai and S. Khorram, "The effects of image misregistration on the accuracy of remotely sensed change detection," IEEE Trans. Geosci. Remote *Sens.*, vol. 36, no. 5, pp. 1566–1577, Sep. 1998.

[27] L. Bruzzone and D. F. Prieto, "Automatic analysis of the difference image for unsupervised change detection," IEEE Trans. Geosci.Remote Sens., vol. 38, no. 3, pp. 1171–1182, May 2000.

[28] A. Singh, "Digital change detection techniques using remotely-sensed data," Int. J. Remote Sens., vol. 10, no. 6, pp. 989–1003, 1989.

[29] T. Celik, "Unsupervised change detection in satellite images using principal component analysis and *k*-means clustering," *IEEE Geosci. Remote Sens.Lett.*, vol. 6, no. 4, pp. 772–776, Oct. 2009.

[30] T.Celik, ''Multiscale change detection in multi temporal satellite images," *IEEE Geosci. Remote Sens.Lett.*, vol. 6, no. 4, pp. 822–824, Oct. 2009.

[31] X.Zhang, L.Wang, L.C.Jiao, ''An Unsupervised change detection based on clustering combined with multiscale and region growing,".

[32] P.Du, S.Liu, P.Gamba, K.Tan and J.Xia, ''Fusion of Difference Images for Change Detection Over Urban Areas," IEEE .J.Appl.Earth.Observ and Remote Sens.,vol .5,no.4,pp.1076-1085,August 2012.

[33] M.Gong, Z.Zhou, J.Ma,''Change Detection In SAR Images Based On Image Fusion And Improved Fuzzy Clustering,"*IEEETrans.Imag.Process*,vol.21,no.4,pp.2141-2151,April 2012.

[34] A. Rosenfeld and P. De la Torre, ''Histogram concavity analysis as an aid in threshold selection,'' *IEEE Trans. Syst. Man Cybern.* SMC-13,231–235 ~1983.

[35] M. I. Sezan, ''A peak detection algorithm and it's application to histogram-based image data reduction,'' *Graph. Models Image Process.*29, 47–59 ~1985.

[36] N. Ramesh, J. H. Yoo, and I. K. Sethi, ''Thresholding based on histogram approximation,'' IEE Proc. Vision Image Signal Process.142~5!, 271–279 ~1995.

[37] D. E. Lloyd, ''Automatic target classification using moment invariant of image shapes,'' Technical Report, RAE IDN AW126, Farnborough, UK ~Dec. 1985.

[38] J. Kittler and J. Illingworth, ''Minimum error thresholding,'' *Pattern Recogn.* 19.41-47~1986

[39] S. Cho, R. Haralick, and S. Yi, ''Improvement of Kittler and Illingworths's minimum error thresholding,'' *Pattern Recogn.* 22, 609–617~1989

[40] T. Pun, ''A new method for gray-level picture threshold using the entropy of the histogram,'' *Signal Process.* 2(3), 223–237 ~1980.

[41] C. H. Li and C. K. Lee, ''Minimum cross-entropy thresholding,'' *Pattern Recogn.* 26, 617–625 ~1993.

[42] H. D. Cheng, Y. H. Chen, and Y. Sun, ''A novel fuzzy entropy approach to image enhancement and thresholding,'' *Signal Process*. 75,*277–301 ~1999.*

[43] J. N. Kapur, P. K. Sahoo, and A. K. C. Wong, ''A new method for gray-level picture thresholding using the entropy of the histogram,''*Graph. Models Image Process.* 29, 273–285 ~1985.

[44] J. C. Yen, F. J. Chang, and S. Chang, ''A new criterion for automatic multilevel thresholding,'' *IEEE Trans. Image Process.* IP-4, 370-378 ~1995.

[45] L. Hertz and R. W. Schafer, ''Multilevel thresholding using edge matching,'' *Comput. Vis. Graph. Image Process.* 44, 279–295 ~1988.

[46] A. Rosenfeld, ''The fuzzy geometry of image subsets,'' Pattern Recogn.Lett. 2, 311–317 ~1984.

[47] W. H. Tsai, ''Moment-preserving thresholding: A new approach,''Graph. Models Image Process. 19, 377–393 ~1985.

[48] Y. Liu and S. N. Srihari, ''Document image binarization based on texture analysis,'' *Proc. SPIE* 2181, 254–263 ~1994.

[49] Y. Liu, R. Fenrich, and S. N. Srihari, ''An object attribute thresholding algorithm for document image binarization,'' ICDAR'93: Proc. 2nd Intl. Conf. Document Anal. Recog., pp. 278–281 ~1993.

[50] K. Ramar, S. Arunigam S. N. Sivanandam, L. Ganesan, and D. Manimegalai, ''Quantitative fuzzy measures for threshold selection,'' *Pattern Recogn. Lett.* 21, 1–7 ~2000.

[51] X. Fernandez, ''Implicit model oriented optimal thresholding using Kolmogorov-Smirnov similarity measure,'' ICPR'2000: Intl. Conf.Patt. Recog., pp. 466–469, Barcelona ~2000.

[52] Y. Nakagawa and A. Rosenfeld, ''Some experiments on variable thresholding," *Pattern Recogn.* 11(3), 191–204 ~1979

[53] W. Niblack, An Introduction to Image Processing, pp. 115–116, Prentice-Hall, Englewood Cliffs, NJ ~1986.

[54] J. M. White and G. D. Rohrer, ''Image thresholding for optical character recognition and other applications requiring character image extraction,'' *IBM J. Res. Dev.* 27(4), 400–411 ~1983.

[55] Y. Yasuda, M. Dubois, and T. S. Huang, ''Data compression for check processing machines,'' *Proc. IEEE* 68, 874–885 ~1980.

[56] L. Bruzzone and D. F. Prieto, "Automatic analysis of the difference image for unsupervised change detection," *IEEE Trans. Geosci. Remote Sens.,*vol. 38, no. 3, pp. 1171–1182, May 2000.

[57] Jun zhang,and Jinglu Hu.Image Segmentation Based on 2D Otsu Method with Histogram Analysis[C]/proc of International Conference on Computer Science and Software Engineering, 2008:105-108.

[58] N. Otsu, ''A threshold selection method from gray level histograms,'' *IEEE Trans. Syst. Man Cybern.* SMC-9, 62–66 ~1979.

[59] T. Celik, "Unsupervised change detection in satellite images using principal component analysis and *k*-means

clustering," *IEEE Geosci. Remote Sens.Lett.*, vol. 6, no. 4, pp. 772–776, Oct. 2009.

[60] Volpi.M, Tuia.D, Campo-Valls.G and Kanevski.M, "Unsupervised Change Detection with kernels",IEEE Trans. *Geosci. Remote Sens.*,vol.9, no. 6, pp. 1026–1030, 2012.

[61] J. Bezdek, Pattern Recognition With Fuzzy Objective Function Algorithms.New York: Plenum, 1981.

[62] D. Pham, "An adaptive fuzzy C-means algorithm for image segmentation in the presence of intensity inhomogeneities," *Pattern Recognit. Lett., vol. 20, pp. 57–68, 1999.*

[63] J. W. Berger, T. R. Patel, D. S. Shin, J. R. Piltz, and R. A. Stone, "Computerized stereochronoscopy and alteration flicker to detect optic nerve head contour change," Ophthalmol., vol. 107, no. 7, Jul. 2000.

[64] H.Liu and Q.Zhou ''Accuracy analysis of remote sensing change detection by rule-based rationality evaluation with post-classification comparison"International Journal of Remote Sensing,25(5),pp.-1037-1050.

[65] P. Rosin, "Thresholding for change detection," Comput. Vis. Image Understanding,vol. 86, no. 2, pp. 79–95, May 2002.

[66] G. H. Rosenfield and A. Fitzpatrick-Lins, "A coefficient of agreement as a measure of thematic classification accuracy," Photogramm. Eng.Remote Sens., vol. 52, no. 2,pp.223–227,1986.