



Using Change Detection and Classification Methods to Study the Al-ahwaz Marshals using Landsat-7 ETM Images

Alyaa Hussein Ali
Physical Department/ College of
Science for women
University of Baghdad
Iraq , Bagdad
aliahusain@ymail.com

Maysaa Raba Naeemah
Physical Department/ College of
Science for women
University of Baghdad
Iraq , Bagdad
ph.maysaa@gmail.com

Israa Jameel Muhsin
Physical Department/ College of
Science
University of Baghdad
Iraq , Bagdad
israaphysics@gmail.com

Abstract: In this paper, six bands data sets taken from the landsat-7 ETM (Enhanced Thematic Mapper) to determine the changes which occur in Al-Ahwaz marshals during 2000 and 2013. Different methods were used to study the changes which occur in this region the subtraction method which Image Difference (ID) and the (PCA) method which is Principal Component Analysis. Campaign the (PCA) with (ID) to produce the best way to detect the changes. The unsupervised classification process which is the modified (K-mean) is used to classify the changes occur in the marshals region and the area surrounding it. This type of classification depends on the color and distance between the classes, a hybrid method which is the (ID) with the K-mean and PCA of (ID) with the K-mean were implemented to produce the best change detection regions.

Keywords: PCA, Image Subtraction, Modified K-mean

I. INTRODUCTION

Remote sensing can be defined as the collection and interpretation of information about an object without being in physical contact with the object. Aircraft and satellites are the common platforms for remote sensing of the earth and its natural resources. A sensor is a device that measures and records electromagnetic energy. Sensors can be divided into two groups. Passive sensors depend on an external source of energy, usually the sun. The most common passive sensor is the photographic camera. Active sensors have their own source of energy, an example would be a radar gun. These sensors send out a signal and measure the amount reflected back. Active sensors are more controlled because they do not depend upon varying illumination conditions [1]. Remote sensing can generally be used to study the change detection which occurs during different times. There are different methods which can be used to monitor the changes. The most common one is the PCA and the traditional methods. Also, the classification process gives good indication about the changes which occur during the years. The best unsupervised classification method is the K-mean.

II. CHANGE DETECTION METHODS

Change detection methods were used to detect and identify the temporal changes in the scene so that the selection of the appropriate method takes on considerable significance for detecting the changes. The technique of PCA produces good estimates of change. Most studies using PCA for change

detection tried to attach physical interpretation, it represents the most empirical approach to mapping change [2].

The digital image change detection can be performed by several methods [3].

- 1- Image differencing
- 2- Image addition
- 3- Image rationing
- 4- Image multiplication
- 5- Principal component analysis (PCA)

The K-means algorithm is one of the simplest and most efficient unsupervised learning algorithms to solve clustering problems in image segmentation. K-means algorithm tends to perform better for homogeneous surface/object as desert area [4][5].

III. ARITHMETIC OPERATIONS

Arithmetic operations are performed on a pixel by pixel basis between two or more images. The actual mechanics of implementing arithmetic operations can be done sequentially, one pixel at a time, or in parallel, where all operations are performed simultaneously. There are four arithmetic operations (addition, subtraction, multiplication and division) that can be applied on an image at the same coordinate system [6].

A. Image Subtraction

Image subtraction is probably the most widely applied change detection algorithm for a variety of geographical environments. It involves subtracting one data of imagery from a second data that has been precisely registered to the first with perfect data. This would result in a data set in which positive

and negative values representing the change areas and zero value represented no change [2]. This operation is also performed on a pixel by pixel basis. In this technique, images of the same area, obtained from times t_1 and t_2 , are subtracted pixel wise. Mathematically, the difference image which can be represented as follows [7].

$$I_d(x, y) = I_1(x, y) - I_2(x, y). \quad \dots\dots\dots(1)$$

Where I_1 and I_2 are the images obtained from t_1 and t_2 , (x, y) are the coordinates of the pixels. The resulting image, I_d , represents the intensity difference of I_1 from I_2 . This technique works only if images are registered.

B. Principal Component Analysis

Principal Component Analysis (PCA) is the general name for a technique which uses mathematical principles to transform a number of possibly correlated variables into a smaller number of variables called principal components. Principal Component Analysis (PCA) is the transformation of the multivariate data to a new set of components where data variation can be expressed by a first few components. PCA achieves this by removing the redundancy in the data set. This redundancy is quantified by the correlation of the variables. Hence, PCA transforms a correlated set of data to an uncorrelated set [8][7].

The mean position of the pixels in the space is defined by the expected value of the pixel vector x , according to equation (2)

$$m = E\{x\} = \frac{1}{k} \sum_{k=1}^k x_k \quad \dots\dots\dots(2)$$

Where (m) is the mean pixel vector and the x_k are the individual pixel vectors of total number k ; E is the expectation operator. While the mean vector is useful to define the average or expected position of the pixels in multispectral vector space, it is the value to have available means by which their scatter or spread is described. This is the role of the covariance matrix which is defined as covariance matrix [9].

The covariance matrix takes the form of correlation matrix whose elements represented the covariance between the images. The diagonal elements of covariance matrix are the variance of each element.

$$\Sigma_x = \begin{bmatrix} c_{11} & c_{12} & c_{13} & \dots & c_{1j} \\ c_{21} & c_{22} & c_{23} & \dots & c_{2j} \\ c_{31} & c_{32} & c_{33} & \dots & c_{3j} \\ \dots & \dots & \dots & \dots & \dots \\ c_{i1} & c_{i2} & c_{i3} & \dots & c_{ij} \end{bmatrix} \quad \dots\dots\dots (3)$$

The eigenvector and eigenvalue of the covariance matrix denoted by a_i and λ_i respectively where $i = 1, 2, 3, \dots, N \times N$ where N is the dimension of the image.

The transformation matrix consists of the eigenvector of the covariance matrix.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1N} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2N} \\ a_{31} & \dots & \dots & \dots & a_{3N} \\ \dots & \dots & \dots & \dots & \dots \\ a_{N1} & a_{N2} & a_{N3} & \dots & a_{N^2} \end{bmatrix} \quad \dots\dots\dots (4)$$

Where A is unity matrix such that $A^{-1} = A^T$.

To compute the principal component the covariance matrix should be diagonalized.

$$A \Sigma_x A^T = \begin{bmatrix} \lambda & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ \vdots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_{N^2} \end{bmatrix} \quad \dots\dots\dots (5)$$

The eigenvalue are $\lambda_1 \dots \dots \lambda_N$ which are uncorrelated [10][2].

PCA has traditionally been used in remote sensing as a means of data compaction. For a typical multispectral image band set, it is common to find that the first two or three components are able to explain virtually all of the original variability in reflectance values. Later components thus tend to be dominated by noise effects. By rejecting these later components, the volume of data is reduced with no appreciable loss of information. Given that the later components are dominated by noise, it is also possible to use PCA as a noise removal technique [11].

IV. THE STUDY AREA

Maysan is one of Iraq's eighteen provinces, it lies in southeastern section of Iraq, within the precipitation valley. The total area of the province is 16072 square kilometers representing 3.7% of the total area of Iraq. The total population in the province of Maysan is about 920,315 people living and resettle differentiated between rural and urban, and it is comprised of 6 districts and 9 regions. This province is bordered from the north by the Wasit Province, from the east by Iran, from the west by Dhi Qar province, and from the south by Basra province figure (1-4). Regarding its astronomical location, it lies between latitudes $35^{\circ}15' - 32^{\circ}45'$ north and longitudes $46^{\circ}30' - 47^{\circ}30'$ east [12].



Figure (1) Maysan Province

The geographical nature of the province is divided between land and water, where the marshlands form more than 40% of the total area extending between its southern and southeastern parts as well as the southwestern part. In addition to that, Degla River flows in the middle of the province, with its numerous branches such as Al-Magar River, al-Mashrah, and Al-Amara, in addition to many branches originating from the eastern side of the province. As for the land areas of the province, they are characterized by being flat in most part, except for some of the eastern areas bordering Iran, which are characterized by the presence of hills that are considered an extension to the Hamreen Mountain Series.

With respect to the features of the surface, it includes several forms, the most important of which are the Degla River banks and its high streams, the low riverbed areas, as well as the marshlands and swamp areas that were dried up during the previous regime, but efforts now continue to rehabilitate them once again. There are also sand dune hills on the western and northeastern sides, some are fixed and others mobile. There is also a series of high hills that extends along the international

borders with the Islamic Republic of Iran, and whose height exceeds (135m) above sea level in some areas. Ripples of valleys lie at the bases of those hills.

The province climate also the annual rate of maximum temperature is 31.5 degrees centigrade, and the minimum temperature is 17.1 degrees centigrade, With respect to rain, the total annual rainfall is 161mm, falling during the period from October till March, and rarely falls during the remaining months of the year . and various types of winds that blow over Maysan[12].

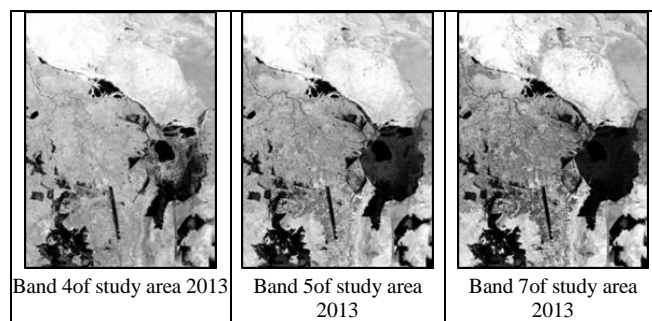
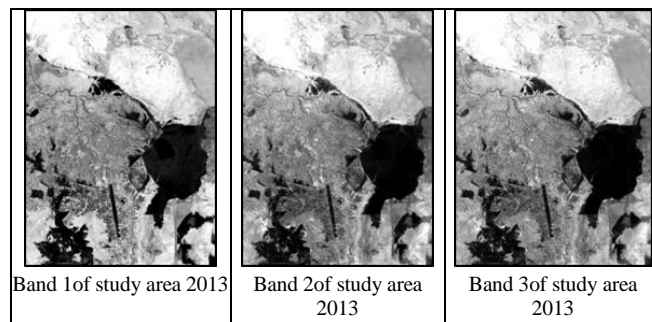
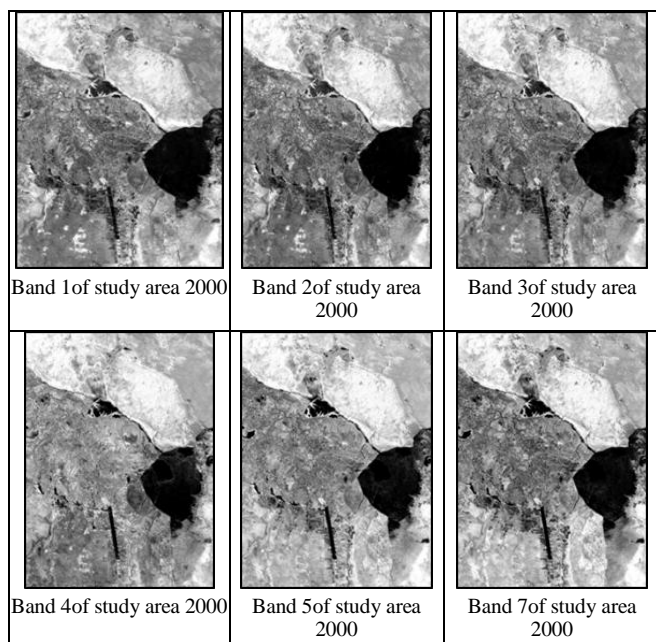
V. METHODOLOGY

Our search focused on detecting the changes occur during (2000 and 2013). Six bands have been taken for the south of Iraq. Table (1) represented the bands information, the first method which is used to detect the changes is the subtraction, figure (2) shows the six bands images for southeastern section of Iraq for two different time, figure (3) shows the image difference between 2000 and 2013. The two images must precisely registered, figure (4) shows the PCA of the images shown in figure (3) the image differences

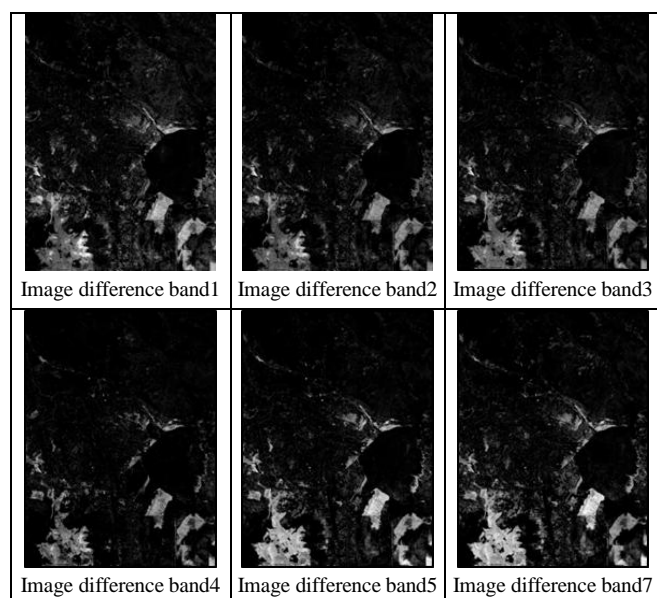
Table (1) application of the Landsat-7(Enhanced Thematic Mapper) [13].

TM Bands

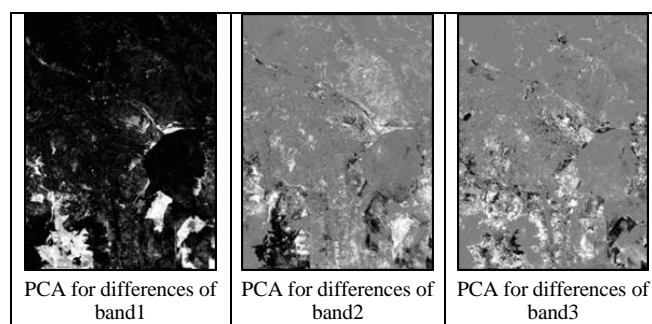
Channel	Wavelength Range (μm)	Application
TM 1	0.45 - 0.52 (blue)	soil/vegetation discrimination; bathymetry/coastal mapping; cultural/urban feature identification
TM 2	0.52 - 0.60 (green)	green vegetation mapping (measures reflectance peak); cultural/urban feature identification
TM 3	0.63 - 0.69 (red)	vegetated vs. non-vegetated and plant species discrimination (plant chlorophyll absorption); cultural/urban feature identification
TM 4	0.76 - 0.90 (near IR)	identification of plant/vegetation types, health, and biomass content; water body delineation; soil moisture
TM 5	1.55 - 1.75 (short wave IR)	sensitive to moisture in soil and vegetation; discriminating snow and cloud-covered areas
TM 6	10.4 - 12.5 (thermal IR)	vegetation stress and soil moisture discrimination related to thermal radiation; thermal mapping (urban, water)
TM 7	2.08 - 2.35 (short wave IR)	discrimination of mineral and rock types; sensitive to vegetation moisture content

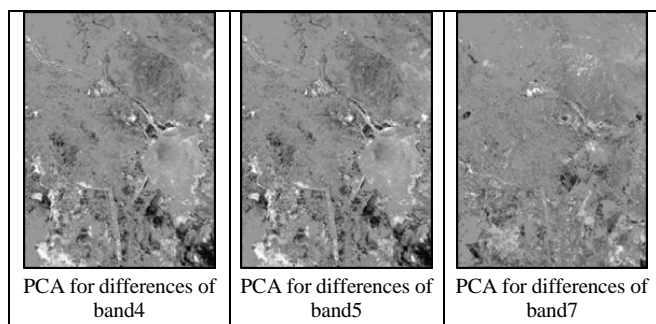


Figure(2) shows the six bands images for southeastern section of Iraq (2000,2013).



Figure(3). Shows image differences (2013-2000).





Figure(4). Shows the PCA of image differences (2013-2000).

VI. MODIFIED K-MEAN

In our study the modified k-mean clustering which depend on the color and distance of the classes has been used. The classes are grouped depend on the color and distance

$$p = \sum_{j=1}^k \sum_{i \in S_j} \text{distance}(x_i^{(j)}, c_j) \quad \dots (6)$$

Where distance is measured between points $x_i^{(j)}$ and the cluster center c_j , S_j is set containing elements of cluster j and k is number of clusters. Algorithm consists of following steps [14].

1. In a set of N points, corresponding to image pixels, choose k points as initial cluster centers (centroids) c_j
2. Assign each point to nearest cluster S_j based on its distance from cluster center c_j .
3. For each cluster S_j compute a mean μ_j of each cluster and set the mean as new cluster center ($c_j = \mu_j$).
4. Repeat the steps 2 and 3 until the centroids no longer move.

VII. CLASSIFICATION OF SUBTRACTED IMAGES

The subtraction is an ordinary method to find the changes occur in the images so, the result of the subtraction process has been adopted to be with the K-means process to produce the real changes occur between 2000 and 2013. Figure (5) shows the classes for each PCA after the (BPD) which we call it (PCD). Figure(6) shows the classes for each band pair difference (BPD). The PCA after the (BPD) has been colored using the density slicing process the classification, show only three classes and the other two classes appear dark, they does not carry any information. as we know the first principal component has the most information so the last PCs have no important information's. Figure(7) represented the graphical distribution of the changed area. Table (2) shows the statistical properties which give indication about the changes during these years.

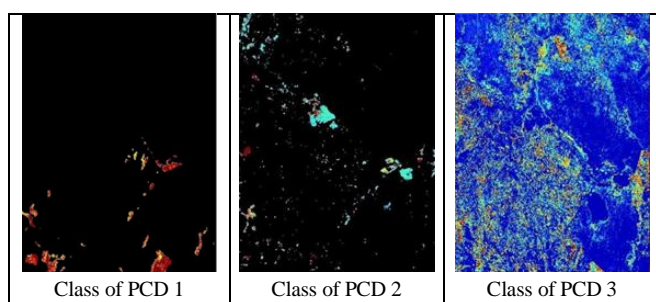
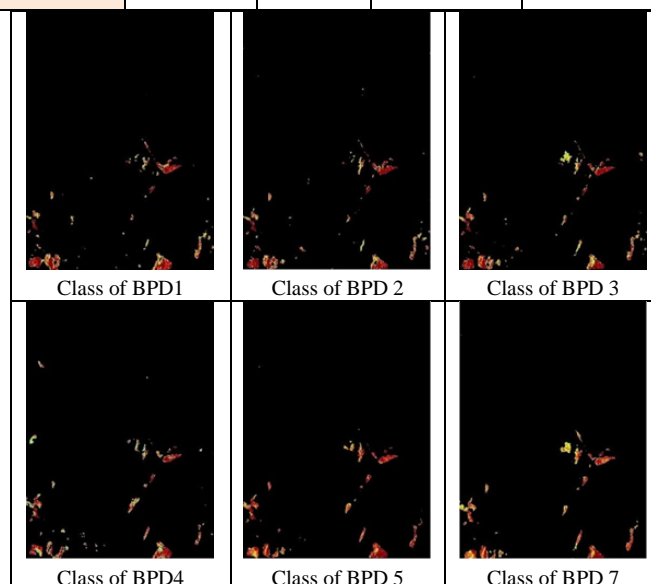


Figure (5) Shows the Classification of PCD

The table (2) shows the statistical properties of the classification of Subtracted Images

Class	Area	mean	Correlation	Stander deviation
PC1	13078	0.0135	0.735	0.0808
PC2	34694	0.0266	0.8931	0.1069
PC3	89994	0.3133	0.7966	0.2044
Class1	11063	0.0106	0.9295	0.0661
Class2	13255	0.0148	0.6637	0.0876
Class3	13678	0.0113	0.9376	0.0691
Class4	13869	0.0137	0.8428	0.0800
Class5	12527	0.0142	0.6593	0.0854
Class6	11745	0.0191	0.6512	0.1065



Figure(6).Shows the classification of the Subtracted Images.

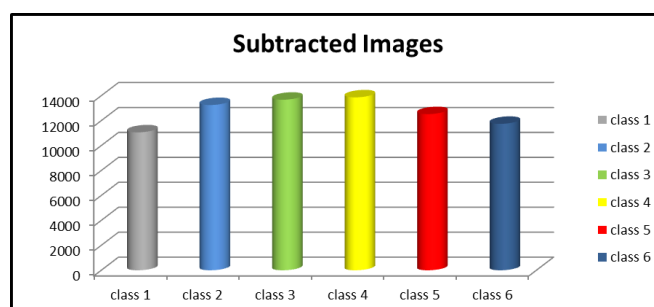


Figure (7) Area of classes of subtracted images using K-Mean classifier technique

VIII. CONCLUSION

In this search, results have been obtained from image differences and principal component image difference(PCD) technique which is very good to detecting the changes occurring in the study area especially in the first principal component and second principal component (PCD1 , PCD2) which contain the most changing information, through practical results it is clear that PCA technique which applied to images in 2000 and in 2013 of the study area directly and then applying the PCA to the process of subtraction, changes highlighted very clearly (the changed area appear clearly). Finally the modified K-mean shows the changed area in classes form, the statistical features for each class has been

calculated to show the changes as can be seen some area increased and other decreased

IX. REFERENCES

- [1] sanderson, R., , 'Introduction to remote sensing 'New mexico state university. Introduction to Remote Sensing and Image Processing.
- [2] Ali, H.A. (1999). Studying the Flooded Area by Principal Component Analysis of Multi-Temporal Landsat Thematic Mapper Data, College of Science , Dep. Physics , Al-Nahrain University, Baghdad, Iraq. (MSc, thesis).
- [3] Celik, T.(2009). Unsupervised Change Detection in Satellite Images Using Principal Component Analysis and k-Means Clustering . IEEE Geoscience
- [4] Hung, C., Purnawan, H., Kuo, B., Letkeman, S.(2009). Multispectral Image Classification Using Rough Set Theory and Particle Swam Optimization, Advances in Geoscience and Remote Sensing, Jedlovec. G.(ed), Intech, ch 28. pp.569-596.
- [5] Lasaponara, R., Masini, N. Satellite Remote Sensing A New Tool for Archaeology, Series: Remote Sensing and Digital Image Processing, Vol. 16, Dordrecht, The Netherlands, Springer, 2012, 364pp.
- [6] Nidhi, Kumari, P. Brain Tumor And Edema Detection Using Matlab7.6.0.324 , International Journal Of Computer Engineering & Technology (IJCET), Volume 5, Issue 3, March (2014), pp. 122-131.
- [7] Ilsever, M., Unsalan, C. Two-Dimensional Change Detection Methods, Remote Sensing Applications. pp.71. Springer, London (2012).
- [8] Russ, J. C. The Image Processing Handbook (Third Edition). CRC Press LLC (1999).
- [9] Richards , J.A, Jia, X. (2006). Remote Sensing Digital Image Analysis: An Introduction, 4th edition : Springer – Verlag, New York, PP 439.
- [10] Gonzalez, R. C. and Woods, R. E. (2008). Digital Image Processing, 3rd ed., Prentice Hall, Upper Saddle River, NJ.
- [11] Eastman, J. R. Idrisi Andes Guide to GIS and Image Processing, , Clark University, Worcester, 2006.
- [12] Republic of Iraq Ministry of Planning. Maysan Provincial Development Strategy. Iraq Information Portal, Web.sep,2011.
- [13] Canada Centre for Remote Sensing. 2007. Tutorial: Fundamentals of Remote Sensing.
http://www.ccrs.nrcan.gc.ca/resource/tutor/fundam/chapter2/08_e.php.
- [14]] Šarić, M., Dujmić, H., & Russo, M. (2013). Scene Text Extraction in HSI Color Space using K means Algorithm and Modified Cylindrical Distance. Przegląd Elektrotechniczny, pp. 117-121.