# Identification of Land and Water Regions in a Satellite Image: A Texture Based Approach

Amandeep Verma
*Assistant Professor in Computer Science,*
*Punjabi University Regional Centre for IT & Mgmt., Mohali*

***Abstract*: - Vision is the most important resource of information for human beings which contain several activities. Amongst these activities, object recognition and classification are widely used. Although, images are representation of vision which can be interpreted by machines, some images often do not exhibit regions of uniform intensity, but these images may contain variation of intensities which form certain repeated patterns. These repeated patterns are known as visual texture. These patterns can be the result of physical surface properties such as roughness, reflectance and color difference of surface. This special variation in pixel intensities is useful in a variety of applications. One such application is analysis of satellite images, whose purpose is to identify various regions in a given image. In this paper we try to exploit knowledge of texture i.e. repeated patterns of non uniform intensity, to recognize regions in the satellite images. There are many approaches used for texture analysis, mainly these are categorized as statistical, structural, filter and model based. In statistical approaches, a GLCM approach is used to find various regions i.e. land and water in a satellite image by extracting its texture information.**

***Keywords*- Texture, Satellite Images, GLCM, Clustering.**

## 1. INTRODUCTION

In many machine vision and image processing algorithms, simplifying assumptions are made about the uniformity of intensities in local image regions. However, images of real objects often do not exhibit regions of uniform intensities [14]. The texture of an image reflects changes of intensity values of pixels which might be contain information of geometric structure of objects so this technique does contain not only the information which is representative for visual characteristics, but also characteristics which can not be visually differentiated. We recognize texture when we see it but it is very difficult to define as there is no precise definition. The"definition" of texture is formulated by different people depending upon the particular application. In the Webster dictionary, texture is defined as"the character of a surface as determined by the arrangement, size, quality, and so on" or "the arrangement of the particles or constituent parts of any material as it affects the appearance or feel of the surface". Thus, the main objective of texture analysis is to extract useful textural information from an image and measures texture directly from the image based on the local spatial variation of intensity or color.

Satellite imagery has large ground coverage's and is provided in digital form than aerial photographs. Therefore, it is more suitable for automated image processing and analysis and is more practical for large areas of interest. Enormous amount of data is being sent by satellites and automation of its analysis is necessary hence we have chosen to analyze satellite images through their texture property.

The paper is organized as follows: In section 2 the literature related to texture analysis is reviewed. In section 3 Gray Level Co-occurrence matrix approach is explained. Section 4 is about the results and discussions and Section 5 list out some directions to improve present work.

## 2. REVIEW OF LITERATURE

Texture analysis methods have been utilized in a variety of application domains ranging from segmentation of satellite images [2] to defect inspection of textile products [13]. These methods are also extensively used in medical applications [10, 6]. The problem of zone classification in a document [15] as text and non-text area, text detection in images [7] can be also addressed by texture features. The classification of terrain [19] on the basis of texture measures is also explored.

With reference to several survey papers [12, 14] texture analysis techniques used for visual inspection are categorized in four ways: statistical approaches, structural approaches, filter based approaches, and model based approaches. Statistical methods [1] analyze the spatial distribution of gray values, by computing local features at each point in the image, and deriving a set of statistics from the distributions of the local features. A large number of statistical texture features have been proposed, ranging from first order statistics to higher order statistics. Amongst many, histogram statistics [18], co-occurrence matrices [20], generalized co-occurrence matrices [17] autocorrelation [16], and local binary patterns [9] have been applied to visual inspection. Structural approaches [18] represent texture by well defined primitives (micro-texture) and a hierarchy of spatial arrangements (macro-texture) of those primitives. Filter based approaches [5] largely share a common characteristic, which is applying filter banks on the image and compute the energy of the filter responses. Model based approaches [14] tries to capture dependencies among neighboring pixel values by assigning an analytical or mathematical function to the textured images. The above mentioned approaches are not mutually exclusive as combined statistical and structural approaches [3] are also used for the purpose of texture classification.

Co-occurrence matrices introduced by [20], GLCM is one of the earliest texture analyzers which is still of interest in many studies. Since the beginning of the 70's many researchers have studied GLCM theory and have practically implemented it in a wide range of texture analysis problems.

A texture analysis technique [11] is proposed for the remote sensing images by the use of a bilinear spatial filter, which is available for extracting the texture difference of the scene objects on the basis of a priori knowledge of their two-dimensional autocorrelations as the spatial signature. Markovian modeling and fuzzy classification [8] is used to extract urban areas from satellite images.

## 3. The Approach

Texture analysis has been a major area of research in past few years. A lot of approaches have been proposed for texture analysis. My aim is to use one of these approaches for satellite images. Satellite images have also received a lot of attention during this time. GLCM also known as Harlick's features are used for this purpose. The present study is constrained to classification of regions of the image into two categories i.e. land and water.

Traditional spectral classification of remote sensing data applied on per pixel basis and they do not take into consideration the useful spatial information available in neighboring pixels. Although spatial information extraction has been greatly explored, there have been limited attempts to enhance classification by combining spectral and spatial information. This study aims to explore the potential of utilizing texture spatial variability using Grey Level Co-occurrence Matrix (GLCM) texture measures [20].

The grey-level co-occurrence matrix (GLCM) is a spatial dependence matrix of relative frequencies in which two neighboring pixels that have certain grey tones and are separated by a given distance and a given angle occur within a moving window [20]. A co-occurrence matrix, also referred to as a co-occurrence distribution, is defined over an image to be the distribution of co-occurring values at a given offset. Mathematically, a co-occurrence matrix C as given in Bayer's tutorial [4] is defined over an $n \times m$ image I, parameterized by an offset ($\Delta x,\Delta y$), as:

$$C(i,j) = \sum_{p=1}^{n}\sum_{q=1}^{m} \begin{cases} 1 \; if \; I(p,q) = i \; and \; I(p+\Delta x, q+\Delta y) = j \\ 0 \; otherwise \end{cases}$$

The 'value' of the image originally referred to the grayscale value of the specified pixel. The value could be anything, from a binary on/off value to 32-bit color and beyond. A 32-bit color will yield a 232x232 co-occurrence matrix. GLCM of an image is computed using a displacement vector d, defined by its radius δ and orientation θ.

Consider a 4x4 image represented by following matrix with four gray-tone values 0 through 3. A generalized GLCM for that image is also shown where #(i,j) stands for number of times gray tones i and j have been neighbors satisfying the condition stated by displacement vector d.

| | | | |
|---|---|---|---|
| 0 | 0 | 1 | 2 |
| 0 | 1 | 0 | 0 |
| 2 | 2 | 3 | 3 |
| 3 | 2 | 1 | 0 |

| Gray Tone | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 0 | #(0,0) | #(0,1) | #(0,2) | #(0,3) |
| 1 | #(1,0) | #(1,1) | #(1,2) | #(1,3) |
| 2 | #(2,0) | #(2,1) | #(2,2) | #(2,3) |
| 3 | #(3,0) | #(3,1) | #(3,2) | #(3,3) |

The four GLCM for angles equal to 0°, 45°, 90° and 135° and radius equal to 1 are shown in figure 1 a-d.

| | | | |
|---|---|---|---|
| 2 | 4 | 0 | 0 |
| 4 | 0 | 2 | 0 |
| 0 | 2 | 2 | 2 |
| 0 | 0 | 2 | 2 |

Figure 1a

| | | | |
|---|---|---|---|
| 2 | 2 | 1 | 2 |
| 2 | 0 | 1 | 1 |
| 1 | 1 | 2 | 0 |
| 2 | 1 | 0 | 0 |

Figure 1b

| | | | |
|---|---|---|---|
| 2 | 2 | 2 | 3 |
| 2 | 0 | 1 | 1 |
| 2 | 1 | 2 | 1 |
| 3 | 1 | 1 | 0 |

Figure 1c

| | | | |
|---|---|---|---|
| 2 | 0 | 2 | 1 |
| 0 | 2 | 1 | 1 |
| 2 | 1 | 0 | 2 |
| 1 | 1 | 2 | 0 |

Figure 1d

These are symmetric matrices hence evaluation of either upper or lower triangle serves the purpose. Frequency normalization can be employed by dividing value in each cell by the total number of pixel pairs possible. Hence the normalization factor for 0° would be (Nx-1) x Ny where Nx represents the width and Ny represents the height of the image. The quantization level is an equally important consideration for determining the co-occurrence texture features. Also, neighboring co-occurrence matrix elements are highly correlated as they are measures of similar image qualities.

The value of δ can be any value ranging from 1 to 10, but δ = 1 or 2 yields better results as a pixel more likely to be correlated to other closely related pixel than the one located far away. Traditionally, GLCM is dimensioned to the number of gray levels G and stores the co-occurrence probabilities Pi. To determine the texture features standard statistics are applied to each GLCM by iterating through the entire matrix. In this study (1, 0) spatial relationship is used i.e. δ =1 and θ=0 east and there would be same GLCM for θ=180.

In process the first step is to find the co-occurrence matrix as demonstrated above. After that in order to make matrix symmetric find the transpose of matrix and add it with the obtained matrix. The rationale behind this is as in the calculation of first step transition from one gray level to other is counted but this number also applicable in other way e.g. if there is a transition from gray level 23 to 47 then this also implies a gray level transition from 47 to 23. After making the GLCM symmetrical, there is still one step to take before texture measures can be calculated. The measures require that each GLCM cell contain not a count, but rather a probability.

It is only an approximation because a true probability would require continuous values, and the grey levels are integer values, so they are discrete. This process is called normalizing the matrix. Normalization involves dividing by the sum of values.

$$P_{i,j} = \frac{V_{i,j}}{\sum_{i,j=0}^{N-1} V_{i,j}}$$

where i is row and j is column of GLCM. Most texture calculations are weighted averages of the normalized GLCM cell contents. A weighted average multiplies each value to be used by a factor (a weight) before summing and dividing by the number of values. The weight is intended to express the relative importance of the value. The texture measures from GLCM broadly divided into three groups namely- Contrast, Orderliness and Descriptive statistics. The contrast group measures use weights related to the distance from the GLCM diagonal. Orderliness means how regular the pixel values within the window. The third group measure consists of statistics derived from GLCM. In this work GLCM mean texture measure using either of the following equation can be used as they both produce same value.

$$\mu_i = \sum_{i,j=0}^{N-1} iP(i,j) \text{ OR } \mu_j = \sum_{i,j=0}^{N-1} jP(i,j)$$

Memory requirement for storing GLCM is very expensive (a 512 x 512 image with 256 gray levels result 256 x 256 x 512 x 512 values). Moreover the resultant matrix is a sparse matrix. Therefore instead of matrix a linked list is used to store the non-zero values with i and j values of gray level. This also improves the computation time. The members of the node are as follows.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| left | ipix | ypix | i | j | count | probability | right |

left – pointer to previous node
ipix,ypix – coordinates of pixel
i,j – neighbors of gray level
count – number of gray levels with these neighbors
probability – occurrence of this gray level transition
right – pointer to next node

The linked list is doubly to make traversing more efficient and even it is sorted one which can be useful for further investigations. The major advantage of this technique is reduction in computational demands as compared to GLCM, although it results in additional computational overhead to sort the list. Texture measures are calculated using this Gray Level linked list.

The final step of this study is to cluster the regions in image on the basis of texture measures. In this work GLCM mean texture measure is used for clustering the various regions. For the purpose of clustering standard kMeans clustering is used. The motive of present work is mainly to distinguish land and water so 2Means clustering is used. For demonstrative purpose 3Means clustering is also done.

The basic steps used for clustering are

*Iterate until stable (= no object move group):*

1. *Determine the centroid coordinate*
2. *Determine the distance of each object to the centroids*
3. *Group the object based on minimum distance*

In summary the steps in the approach are as follows

1. *Decide the window size, spatial displacement and angle of orientation.*
2. *Find the Gray Level Liked List for every pixel of the image by the method s explained above which gives the important measure probability of transition from gray level i to gray level j and in other way around.*
3. *Calculate the various texture measures for analysis purposes*
4. *As all measures are capable of distinguishing the various regions, select the one which is more appropriate for the given image domain (GLCM Mean seems more suitable for the images used in this study)*
5. *Use kMeans Clustering to find out the various clusters ( k =2 in present work )*

The trade-offs involved in GLCM computation is the window size and computation time. It is obvious that larger the size of window better would be the result but it increase the computation time. Increase in the size of window would produce better results only up to some limit in the size of window as larger the window does not produce any significant change in the results. The reason for this as the distance of the neighbor increases, less of its effect on the central pixel.

The results after clustering can be visualized in the following images. All of the input images are gray scale of size 512 x 512. The output/clustered image is with two gray scale values. One of the gray scale value shows the regions in the image which are water and the other show the regions of land. The approach can be extended for more than two classes of regions. This is shown in image (2d) where the image is clustered into three classes.

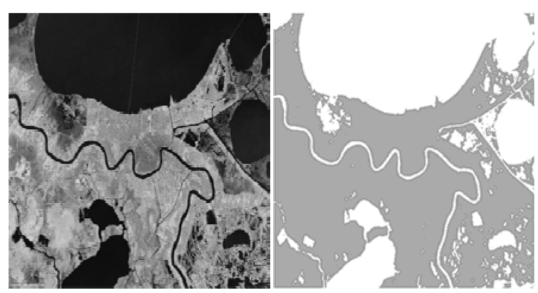

Figure 2a

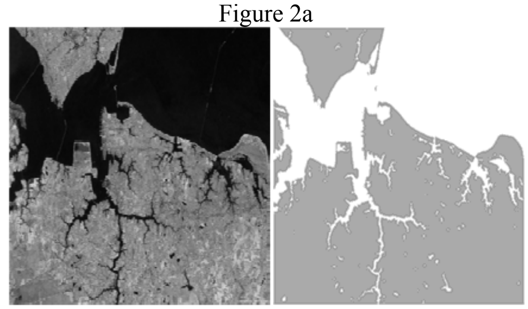

Figure 2b

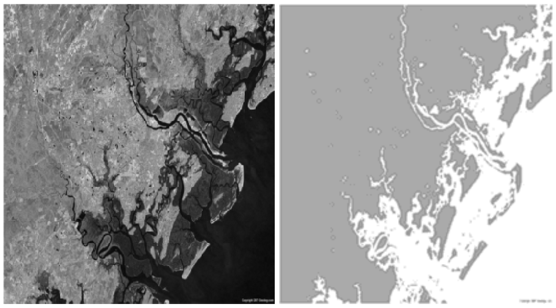

Figure 2c

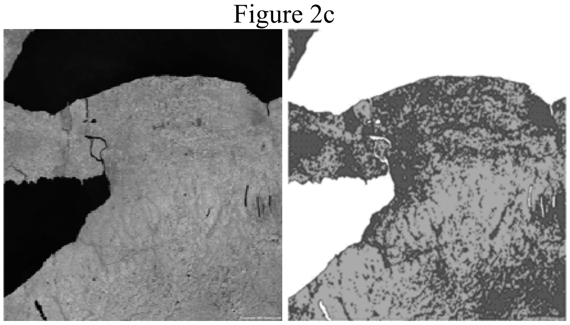

Figure 2d

**Table 1**

| Water Classification | GLCM |
|---|---|
| Image1 | 95.36983 |
| Image 2 | 99.60345 |
| Image 3 | 98.44054 |

**Table 2**

| Land Classification | GLCM |
|---|---|
| Image1 | 99.1776 |
| Image 2 | 99.25626 |
| Image 3 | 99.25568 |

## 4. Results and Discussions

In order to evaluate the performance of the approach the given images are threshold for land and water. The number of pixels representing land and water of threshold image is compared with the pixels in clusters of GLCM output image to find the percentage of correctly classified pixels.

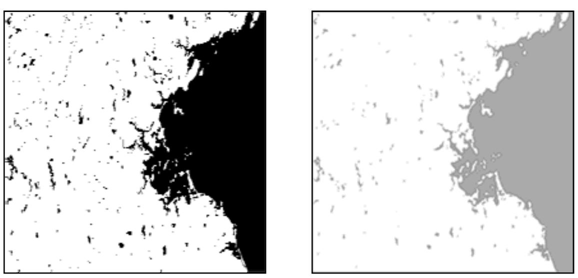

Image 1

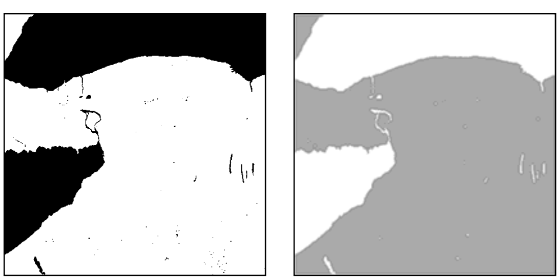

Image 2

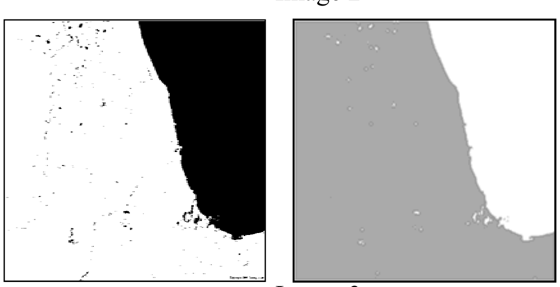

Image 3

The results reveal that the approach used for identification of land and water in a satellite image provides good results. The trade-offs involved in GLCM computation is the window size and computation time. It is obvious that larger the size of window better would be the result but it increase the computation time. Increase in the size of window would produce better results only up to some limit in the size of window as larger the window does not produce any significant change in the results. The reason for this as the distance of the neighbor increases, less of its effect on the central pixel.

## 5. Future Directions

– Efficient implementation of data structure for storage of GLCM.
– Comparison of results with different directions, different distances (i.e. spatial displacement and orientation) and with different window sizes
– Classification of regions into three or more classes.

## References


[1] Srinivasan G.N. and Shobha G., "Statistical Texture Analysis", *Proceedings of World Academy of Science, Engineering and Technology*, vol 36, pp. 1264-1269, 2008.

[2] Tsaneva Mariana, "Texture Features for Segmentation of Satellite Images", *Cybernetics and Information Technologies*, vol 8, No. 3, pp. 73-85, 2008.

[3] Umarani C., Ganesan L. and Radhakrishanan S.,"Combined Statistical and Structural Approach for Unsupervised Texture Classification", *International Journal of Imaging Science and Engineering*, vol 2, No. 1, pp. 1434-1440, 2008.

[4] Bayer Mryka Hall,"GLCM tutorial" ,http://www.fp.ucalgary .ca/mhallbey/tutorial.htm, 2007

[5] Vyas Vibha S. and Rege Priti, "Automated Texture Analysis with Gabor filter", *GVIP Journal*, vol 6 (1), pp. 35-41, 2006.

[6] Iakovidis D. K, Maroulis D.E., Karkanis S.A. and Brokos A.,"A Comparative Study of Texture Features for the Discrimination of Gastric Polyps in Endoscopic Video", *Proceedings of the 18th IEEE Symposium on Computer-Based Medical Systems*, pp. 575 – 580, 2005.

[7] Kwang L. K., Keechul J. and Jin H. K., "Texture-based approach for text detection in images using support vector machines and continuously adaptive mean shift algorithm", *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 25(12), pp. 1631-1639,2003.

[8] Lorette A, Descombes X and Zerubia J, "Texture Analysis through a Markovian Modelling and Fuzzy Classification: Application to Urban Area Extraction from Satellite Images," *International Journal of Computer Vision*, vol 36, pp. 221-236, 2000

[9] Ojala Timo, Pietikainen Matti and Maenpaa Topi, "Gray Scale and Rotation Invariant Texture Classification with Local Binary Patterns",

*Proceedings of European Conference on Computer Vision*, pp. 404 – 420, 2000.

[10] Qiang Ji, Engel John and Craine Eric, "Texture Analysis for Classification of Cervix Lesions", *IEEE Transactions on Medical Imaging*, vol 19, No. 11, pp. 1144-1149, 2000

[11] Koji Takemura, Fuminori Kimura and Koji Hamauka,, "Texture Analysis of Satellite Image by using Spatial filters," *Proceedings of the Annual Conference of the Institute of Systems, Control and Information Engineers*, vol 43, pp 277-278, 1999

[12] Materka A. and Strzelecki M., "Texture Analysis Methods - A Review", *Technical University of Lodz, Institute of Electronics, COST B11 report, Brussels*, 1998.

[13] Ozdemir S., Baykut A., Meylani R., Ercil A. and Ertuzun A.,"Comparative Evaluation of Texture Analysis Algorithms for Defect Inspection of Textile Products", *Proceedings of International Conference on Pattern Recognition*, vol 2, pp. 1738-1740, 1998.

[14] Tuceryan M. and Jain A. K., "Texture Analysis", In Chen C.H., Pau L.F.,Wang P.S.P. (eds),*The Handbook of Pattern Recognition and Computer Vision (2nd Edition), World Scientific Publishing Co*., pp. 207-248, 1998.

[15] Chetveriko Dmitry, Liang Jisheng, Komuves Jozsef and Haralick R. M.,"Zone Classifications using Texture Features", *Proceedings of IEEE Conference on Pattern Recognition*, vol 3, pp. 676-680, 1996.

[16] Kurita T. and Otsu N.,"Texture Classification by Higher Order Local Autocorrelation Features", *Proceedings of Asian Conference on Computer Vision*, pp. 175- 178, 1993.

[17] Davis L. S., Johns S. A. and Aggarwal J. K, "Texture Analysis Using Generalized Co-Occurrence Matrices", *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 1(3), pp. 251-259, 1979

[18] Haralick Robert M., (1979), "Statistical and Structural Approaches to Texture," *Proceedings of the IEEE*, vol 67, No 5, pp 786-804, 1979

[19] Weszka J, Dyer C and Rosenfeld A , "A comparative study of texture measures for terrain classification", *IEEE Transactions on Systems, Man, and Cybernetics*, pp. 269-285, 1976.

[20] Haralick Robert M., Shanmugam K. and Dinstein Itshak,"Textural features for Image Classification", *IEEE Transaction on System, Man and Cybernetics*, vol 3, No 6, pp. 610-618, 1973