Volume 2, No. 10, October 2011
Journal of Global Research in Computer Science
RESEARCH PAPER
Available Online at www.jgrcs.info
ISSN-2229-371X

A NOVEL STATISTICAL APPROACH FOR SEGMENTATION OF SAR IMAGES

Debabrata Samanta[*1], Goutam Sanyal[2]
[*1] Department of CSE, National Institute of Technology, Durgapur ,West Bengal, India - 713209
debabrata.samanta369@gmail.com[1]
[2] Department of CSE, National Institute of Technology, Durgapur, West Bengal, India - 713209
nitgsanyal@gmail.com[2]

Abstract: The statistical parameters include high order image statistics like skewness which describe the shape and symmetry of the image. SAR images data are the result of a coherent imaging system that produces the impair noise fact. Image segmentation is the process of unscrambling or grouping an image into different parts. The good feat of recognition algorithms based on the quality of segmented image. The main problem in SAR image application is correct segmentation. In this paper, we consider the problem of SAR image segmentation by a statistical parameter.

INTRODUCTION

SAR images come across gradually more wide applications, as SAR sensor does not based on weather condition and can infiltrate clouds. SAR image Segmentation is a main examples of low-level operators that providing the basic information for classification of different region. Expected values and variances are important concepts in statistics. They are used to describe distributions, to evaluate the performance of estimators, to obtain test statistics in hypothesis testing, and many other applications.
The motivation of this work is to develop a novel segmentation algorithm, which can be used to segment the SAR images and recover the overall accuracy.
There are different methods of segmenting an image based on the intensity value, including a simple thresholding, quantization of the image, etc. The important factor in this method of segmentation is choosing a threshold for classification of a region to a particular segment. Deciding on a generic threshold that would work for all images is essentially impossible and hence segmenting an image based on just a pre-defined threshold is not a practical technique. Segmentation algorithms for Synthetic Aperture Radar images have been suggested in the field of categorization. C. H. Fosgate, W. W. Irving and W.C.Karl proposed statistical differences in SAR imagery of different terrain types. Lionel Bombrun used heterogeneous clutter models to describe the SAR data. A.A.G., C.L.M. used multi-scale PolSAR data filtering technique based on BPT for SAR image segmentation E.V.D.Lucca and A.C.Frery proposed MUM and RWSEG general means for better segmentation.

PROPOSED METHOD

A. Segmentation based on gray level intensity

Generally the popular method of segmenting an image is based on the gray scale intensity values in the image. Segmenting an image into different regions is done by clustering each gray level intensity value tightly to a mean value or centroid. This is an iterative method, which needs only the number of segments to be hard-coded and adaptively finds the threshold to classify regions within the image.
As an initial approximation, the number of segments is pre-defined as *N*, and as many different, equally spaced mean values are obtained based on the maximum and minimum gray level values in the image, as given in the following equation

$$\mu=(i*p)+min, \text{ ----------------------(1)}$$

Where, μ, is the mean of the i^{th} class,

$$p=(max-min)/(n-1) \text{ ----------------------(2)}$$

max is the maximum gray level intensity value, and min is the minimum gray level intensity value. Each of the pixel values in the image is now assigned to the bin of the corresponding class based on their distance from the mean value of each class. As a measure to identify the distance, the linear distance is calculated as

$$dist = I_{x,y} - \mu \text{ ----------------------(3)}$$

Where $I_{x,y}$, is the pixel intensity at the location *(x,y).*
After all the pixels have been placed into a class, the mean value of all the pixels in a given class is evaluated and used as the new reference mean for assigning pixels in the next iteration. This process is repeated until the mean values converge or a given number of iterations have been completed. Finally, the image is classified into N regions based on the mean values.
The value of N is chosen based on the number of features that require extraction and their corresponding gray level intensities. Features of interest to us in the images are houses, roads, roofs, sidewalks, driveways and trees.

B. Mean

The statistical mean gives us the average value of all the pixels for a particular feature. Mathematically, the statistical mean of a set of variables is expressed as

$$\mu=\sum w_{ij} p \;\; w_{ij} \text{ ------------ 4}$$

The summation is carried out over all the gray level values in the set. It is important to notice that, even though the mean gives a good estimate of the characteristics of a feature, it might not give a good representation of the damage or change in the region.

C. Variance

Variance is the second moment of a matrix and is a measure of gray-level contrast that can be used to establish descriptors of relative smoothness. Mathematically, it is defined as the mean of squares of differences between respective samples (gray-level intensities) and their mean and is expressed as

$$\delta^2=\frac{\sum \mathrm{W_{ij}}-\mu^{2}}{n}------------- 5$$

Where δ^2 the variance or square of the standard deviation, and n is the number of samples in the region. An estimate of the difference in the variance values of the region between the predicator and post-disaster images is an indication of the level of change.

D. Entropy

Entropy is a measure of disorder, or more precisely unpredictability. The probability of a SAR images intensity occurring at particular pixel in 'k', where
'k' is the set of all pixels in a SAR image, is defined as $p\{n\}\log(p\{n\})$. The sum of all of these probability makes the Entropy of 'P', so,

$$E(k)=\sum_{n\in k} p\{n\}\log(p\{n\})$$

Where $p\{n\}$ is the probability mass function of particular pixel in 'k'.

E. Median

The median of a variable is defined as the middlemost value when its values are arranged in ascending or descending order of magnitude. In other hand, the median divides the whole set of binary image values in two parts such that half of the observations are less than or equal to it and half are more than equal to it. So median is,

$$Me = l+\frac{\frac{n}{2}-F_l}{f_{Me}}*C$$ Where f_{Me} and C are the frequency and with class interval, respectively.

F. Mode

The mode of a variable is that values of the variable which has the height frequency or frequency density. A binary image can be fitted with a smooth frequency curve, and then the mode is the abscissa of the height point on the curve. So the mode is ,

$$Mo = X_i+\frac{f_m-f_{m-1}}{2f_m-f_{m-1}-f_{m+1}}*C$$

Where, C is the common with of the class.

G. Skewness

Skewness, or the third standardized moment, is a measure of the asymmetry of the probability distribution of a real valued random variable.

If skewness is negative, the data are spread out more to the left of the mean than to the right. If skewness is positive, the data are spread out more to the right. The skewness of the normal distribution (or any perfectly symmetric distribution) is zero.

First order Skewness: When the frequency distribution is not symmetrical, it depends on first order Skewness. So, first order skewness given below:

First order skewness $=\frac{Men-Mode}{S.D.}$.

Skewness characterizes the degree of asymmetry of a distribution around its mean. The skewness of a distribution is a non-dimensional quantity calculated as

$$Skewness=\frac{1}{N}\sum[\frac{\mathrm{W_{ij}}-\mu^{3}}{\delta^3}]-----(6)$$

Where δ *is* the standard deviation or the square root of the variance. A positive value indicates a deviation of the distribution more towards the brighter side and a negative value towards the darker side.
A distribution is symmetrical if it shows two identical mirror images when it is split down the middle as the normal distribution when a distribution is not symmetrical, it has elongated tail at one side (left or right) and there are more data in this side tail that would be expected in a normal distribution. When it happens at the left, the distribution is negatively skewed or skewed to the left, and the in same way, when it happens at the right the distribution is positively skewed or skewed to the right.

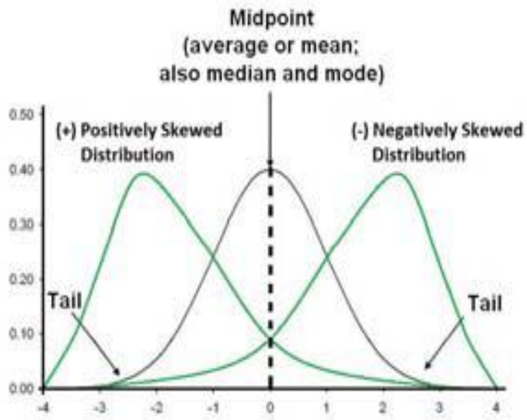

Figure1:- Example of symmetrical, positively skewed and negatively skewed distributions

H. Kurtosis:

The fourth standardized central moment, or kurtosis, is generally defined as a measure reflecting the degree to which a distribution is peaked. In particular, kurtosis provides information regarding the height of the distribution relative to the value of its standard deviation.
The kurtosis of the normal distribution is 3. Distributions that are more outlier-prone than the normal distribution have kurtosis greater than 3; distributions that are less outlier-prone have kurtosis less than 3. The kurtosis of a distribution is defined as

$$K=\frac{E(y-\mu)^4}{\delta^4}--------(7)$$

Where µ is the mean of x, is the standard deviation of y, and E(t) represents the expected value of the quantity t.

The kurtosis value (*K*) can be classified in three general categories according to its derivation from a normal distribution: mesokurtic (normal distribution, *K=0*), leptokurtic (high degree of peakedness, *K>0*) and platykurtic (low degree of peakedness, *K<0*). Figure 2 illustrates these three distribution types.

The kurtogram will provide an indication of, first, the location of eventual subsurface defects, and second, their thermal diffusivity.

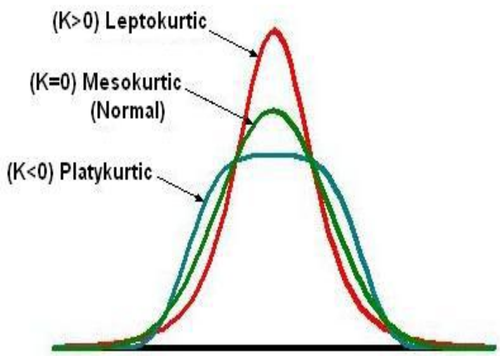

Figure2:- Example of Leptokurtic, mesokurtic and platykurtic distributions.

K-MEANS CLUSTERING ALGORITHM

K-means clustering algorithm is a method of cluster analysis which aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean. K-means clustering Algorithm, known as also C-means clustering, has been applied to variety of areas, including image and speech data compression. The objective function is defined as

$$W=\sum_{j=1}^{M}\sum_{i=1}^{N}|(E_i^{(j)}-V_j)^2|--------(8)$$

Where $(E_i^{(j)}-V_j)^2$ is a chosen distance measure between a data point $E_i^{(j)}$ and the cluster centre Vj is an indicator of the distance of the n data points from their respective cluster centers.

PROPOSED ALGORITHM

- Input: SAR Images of variable size.
- Output: Segmented region of SAR image.

1) Start.
2) Taken a SAR images.
3) Consider a 3X3 window.
4) Calculate the mean, variance SAR Images.
5) Store the color feature in the standardized moment i.e. skewness and fourth standardized central moment i.e. kurtosis.
6) Segmentation is obtained using K-means clustering Algorithm.
7) Stop.

EXPERIMENTAL RESULT

Here SAR images are used to test proposed methodology. At first, the mean, variance was calculated for original image. Images were tested for analyzing window size of 3x3 and 5x5 to find the third standardized moment, fourth standardized central moment. Finally K-means clustering Algorithm was used on Skewness image for segmentation. Now we calculate the features table given below.

Table1: Features table of tested SAR Images.

Features	Image3a	Image4a	Image5a
Mean	106.2200	105.1209	104.6448
Standard deviation	36.6332	34.2693	73.2734
Entropy	6.9004	6.6587	6.2737
Median	103	96	107
Mode	43	37	46
1st order skewness	7	7	6
2nd order skewness	20	20	17
Beta - coefficients	0.0039	0.0043	0.0052
Gama-coefficients	0.0626	0.0658	0.0725
kurtosis	-2.9999	-2.9998	-3

Figure3a
Input SAR image

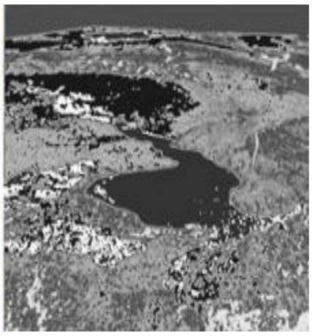

Figure3b
Segmented SAR image

Figure4a
Input SAR image

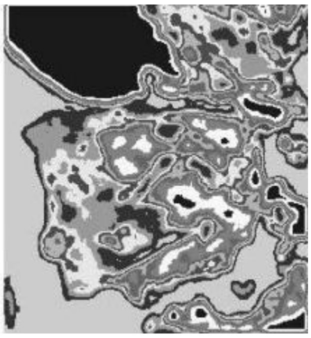

Figure4b
Segmented SAR image

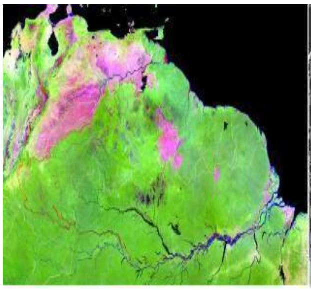
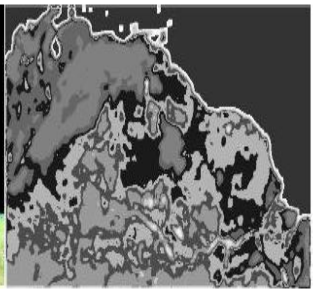

Figure5a
Input SAR image

Figure5b
Segmented SAR image

CONCLUSION AND DISCUSSION

In this paper, we proposed a novel statistical parameter based on skewness and kurtosis for segmentation of SAR image. This technique based on considering a 3X3 window and calculates the mean, variance, mode, median, entropy of that SAR Images. Next, Segmentation is obtained using K-means clustering Algorithm. It was evidenced that this segmentation procedure is a straightforward extension of the filtering algorithm based on standardized moment i.e. skewness. For this reason, our algorithm did not make mistakes; that is, a segmented image very different to get the originality of the SAR images. This may be extended to the color image segmentation. The results from this preliminary study indicated that the proposed strategy was effective.

REFERENCES

[1] Lee J S, Grunes M R, Kwok R. “Classification of multi-look polarimetric SAR imagery based on complex Wishart distribution”, Int. J. Remote Sensing, 1994, 15(11): 2299-2311.

[2] Bischof H, Schneider W, Pinz A J.”Multispectral classification of landsat-images using neural networks ”, IEEE Trans. Geosci. Remote Sensing, 1992, 30(3): 482-490.

[3] Chen C T, Chen K S, Lee J S. “The use of fully polarimetric information for the fuzzy neural classification of SAR images”, IEEE Trans. Geosci. Remote Sensing, 2003, 41(9 Part I): 2089-2100.

[4] Du L, Lee J S. “Fuzzy classification of earth terrain covers using complex polarimetric SAR data”, International Journal of Remote Sensing, 1996, 17(4): 809-826.

[5] Fukuda S, Hirosawa H. “A wavelet-based texture feature set applied to classification of multi frequency polarimetric SAR images”. IEEE Trans. Geosci. Remote Sensing, 1999, 37(5): 2282-2286.

[6] Wenbing Tao,Hai Jin,Liman Liu; Object segmentation using ant colony optimization algorithm and fuzzy entropy, Pattern Recognition Letters 28 (2007) 788796.

[7] H. B. Kekre, Saylee Gharge, Tanuja K. Sarode; SAR Image Segmentation using Vector Quantization Technique on Entropy Images, (IJCSIS) International Journal of Computer Science and Information Security, Vol. 7, No. 3, March 2010.

[8] Debabrata Samanta, Mousumi Paul and Goutam Sanyal, ”Segmentation Technique of SAR Imagery using Entropy*”*, Int. J. Comp. Tech. Appl., Vol 2 (5), 1548-1551, ISSN: 2229-6093, 2011.

[9] A. Nakib, H. Oulhadj and P. Siarry ; Microscopic image segmentation with two-dimensional exponential entropy based on hybrid microcanonical annealing, MVA2007 IAPR Conference on Machine Vision Applications, May 16-18, 2007, Tokyo,JAPAN.

[10] R.L.Euban,” Spline Smoothing and Nonparametric Regression”,Marcel Dekker,New York.

[11] *Debabrata Samanta* and Goutam Sanyal, “Development of Edge Detection Technique for Images using Adaptive Thresholding”, Fifth International Conference on Information Processing (ICIP-2011), CCIS 157, pp. 671-676, 5-7 Aug.2011. @ *Springer*-Verlag Berlin Heidelberg 2011.

[12] H. B. Kekre, Saylee Gharge, Tanuja K. Sarode; SAR Image Segmentation using Vector Quantization Technique on Entropy Images, (IJCSIS) International Journal of Computer Science and Information Security, Vol. 7, No. 3, March 2010.

[13] Chen C T, Chen K S, Lee J S. “The use of fully polarimetric information for the fuzzy neural classification of SAR images”, IEEE Trans. Geosci. Remote Sensing, 2003, 41(9 Part I): pp 2089-2100.

[14] Debabrata Samanta, and Goutam Sanyal*,”* SAR image segmentation using Color space clustering and Watersheds*”*, International Journal of Engineering Research and Applications, Vol. 1, Issue 3, pp.997-999 ,2011.

[15] A.P.Reeves,A Parallel Mesh Moment Computer, Proceedings of the 6th International Conference on Pattern Recognition .October ,pp 465-467.

SHORT BIODATA OF ALL THE AUTHOR

Debabrata Samanta is a member of the IAENG. He obtained his B.Sc. (Physics Honors) in the year 2007, from the Vivekananda Collage, Takurpukur, under Calcutta University; Kolkata, India .He obtained his MCA in the year 2010, from the Academy Of Technology, under WBUT. He is working his PhD in Computer Science and Engg. in the year 2010 from National Institute of Technology, Durgapur, India in the area of Image Processing .He is presently working as a Lecturer of CSE in Abacus Institute of Engg. And Management, West Bengal, India. His areas of interest are Artificial Intelligence, Natural Language Processing and Image Processing. He has published 10 papers in International Journals / Conferences.

Gautam Sanyal is a member of the IEEE. He has received his B.E and M.Tech degree from National Institute of Technology (NIT), Durgapur, India. He has received Ph.D. (Engg.) from Jadavpur University, Kolkata, India, in the area of Robot Vision. He possesses an experience of more than 25 years in the field of teaching and research. He has published nearly 65 papers in International and

National Journals / Conferences. Two Ph.Ds (Engg) have already been awarded under his guidance. At present he is guiding six Ph.Ds scholars in the field of steganography, Cellular Network, High Performance Computing and Computer Vision. He has guided over 10 PG and 100 UG thesis. His research interests include Natural Language Processing, Stochastic modeling of network traffic, High Performance Computing, Computer Vision. He is presently working as a Professor in the department of Computer Science and Engineering and also holding the post of Dean (Students' Welfare) at National Institute of Technology, Durgapur, India.