Automatic Cloud Detection Based on Neutrosophic Set in Satellite Images

Jeethu Mary Mathew
Department of Computer Science
University of Kerala, Kariavattom,
Thiruvananthapuram, India.
jeethu_mm@yahoo.co.in

Surya S.R
Department of Computer Science
University of Kerala, Kariavattom,
Thiruvananthapuram,India.
sooryasr07@gmail.com

Philomina Simon
Department of Computer Science
University of Kerala, Kariavattom,
Thiruvananthapuram, India.
philomina.simon@gmail.com

Abstract— In this paper, an approach for automatic cloud detection and localization in satellite remote sensing images is introduced. Cloud detection is useful in improving the accuracy of land cover classification in cloudy satellite images. The accurate detection of clouds in satellite images is vital for many atmospheric and terrestrial applications. In this paper we propose an algorithm for automatic cloud detection based on neutrosophic set and wavelet transform. The proposed approach uses both color and texture features for cloud detection. The input image is transformed into Lab color model for extracting the color features and gray scale image for extracting the texture features. Transformed images are converted into neutrosophic domain. An indeterminacy reduction operation is performed for getting better results. Finally a Fuzzy C-means clustering is performed on the true subsets. This gives the cloud detected image. This method is efficient in detecting thick clouds and thin clouds in Landsat images. Result analysis shows that the proposed algorithm can effectively detect the thin cloud. The proposed algorithm gives accurate results in less time complexity.

Keywords—*Cloud detection; Neutrosophic Set; Color transformation; Fuzzy Clustering; Landsat ETM+*

I. INTRODUCTION

When the world's first satellite pictures of the atmosphere were viewed, the most remarkable feature was the extensive cloud cover over large parts of earth. Even today, the first feature that one usually notices in the satellite image is the presence of clouds. Persistent cloud covers over many regions cause difficulties in remote sensing with optical satellite imagery. Globally, land scenes are, on average, about 35% cloud covered, as reported by Ju and Roy [1], indicating that cloud covers are usually present in optical satellite images. This phenomenon limits the usage of optical images and increases the difficulty in image analysis. Cloud detection is useful in improving the accuracy of land cover classification in multispectral images. Clouds may seem random in shape and distribution. Cloud cover is a big challenge in optical remote sensing of the earth surfaces, especially over the humid tropical regions. Clouds are generally characterized by higher reflectance and lower temperature than the underlying earth surface. The brightness of clouds in a satellite image is one of the best hints to use. High brightness values in visible satellite images are associated with thick clouds, which tend to reflect much of the sun's light. Therefore thick clouds appear white in visible satellite images. Thin clouds appear darker or even are transparent in visible satellite imagery. Automated cloud detection methods use spatial analysis methods to detect the contrast between reflected energy from clouds and nearby scenes and can be used to determine the extent of cloud cover over a region.

The rest of the paper is organized as follows. Section II describes literature review. Section III explains about neutrosophic set. Section IV discusses the proposed cloud detection algorithm. Section V demonstrates experimental results and analysis. Conclusion is given in Section VI.

II. LITERATURE REVIEW

Cloud detection is a one of the most key processes in weather prediction and regional planning. Detection of clouds helps to improve the efficiency in study of climatic changes. Wang et. al. [2] proposed a scheme to remove clouds and their shadows from remotely sensed images of Landsat TM. The cloud regions can be detected on the basis of the reflectance differences with the other regions. An efficient and robust neural network-based scheme is introduced by Azimi-Sadjadi [3] to perform automatic cloud detection and classification. This method uses an unsupervised Kohonen neural network to classify the cloud contents of an 8×8 blocks in an image into ten different cloud classes. Inputs to the network consist of texture features of each blocks extracted using wavelet transform (WT). Salvador et al. [4] adopted pseudo-Wigner distribution (PWD) to detect the cloud. An automated cloud cover assessment (ACCA) algorithm [5] has been developed for estimating the percentage of cloud cover in Landsat 7 Enhanced Thematic Mapper Plus (ETM+) images. Jiao et al. [6] proposed a Data Ratio (IRDR) algorithm based an intensity threshold and the method of maximum likelihood estimation to classify cloud-covered and cloud-free area. The cloud detection algorithm proposed by Yu Tong et al. [7] detects clouds and shadows in remote sensed images based on MSER (Maximally Stable Extremal Regions). Huang et al.[8] uses clear view forest pixels as a reference to define cloud boundaries for separating cloud from clear view surfaces in a spectral–temperature space for detecting clouds. Since this algorithm is proposed for forest change analysis, it requires forest pixels to determine the cloud boundaries, this method can be applied only to images where forest pixels exist. Neural network-based cloud classification on satellite imagery using texture features

is developed by Bin Tian[9]. In that method, several image transformation schemes such as wavelet transform (WT) and singular value decomposition (SVD) are used to extract the salient texture feature of the data and is then compared with those of the well-known gray-level co-occurrence matrix (GLCM) approach. Cheng and Y. Guo [10] proposed an approach for image segmentation based on neutrosophic set. But, this method works only for gray scale images.

In this paper, a new approach is proposed for cloud detection using neutrosophic set and wavelet transform. Neutrosophic set provides a powerful tool to deal with the indeterminacy of the image. The clouds can be detected effectively by minimizing the indeterminacy. This method is applicable to detect thin and thick clouds in Landsat images. The goal of the proposed work is to automatically detect clouds using color and texture features in Landsat images. The proposed algorithm is efficient and detects most of the clouds in Landsat images. The proposed algorithm gives accurate detection of clouds in Landsat images.

III. NEUTROSOPHIC SET

Neutrosophic set (NS), is proposed by Florentin Smarandache as a new branch of philosophy dealing with the origin, nature and scope of neutralities, as well as their interactions with different ideational spectra [11]. In neutrosophy theory, every event has not only a certain degree of truth, but also a falsity degree and an indeterminacy degree that have to be considered independently from each other. Thus, a theory, event, concept, or entity, {A} is considered with its opposite {Anti-A} and the neutrality {Neut-A}. {Neut-

A} is neither {A} nor {Anti-A}. The {Neut-A} and {Anti-A} are referred to as {Non-A}. According to this theory, every idea {A} is neutralized and balanced by {Anti-A} and {Non-A} [10].

Define T, I, and F as neutrosophic components to represent {A}, {Neut-A}, and {Anti-A}. Let T, I, and F are real standard or non-standard sets of $]^{-}0, 1^{+}[$ with $supT = t\_sup, infT = t\_inf, supI = i\_sup, infI = i\_inf, supF = f\_sup, infF = f\_inf$ and $n\_sup = t\_sup + i\_sup + f\_sup, n\_inf = t\_inf + i\_inf + f\_inf$ [10]. $T, I,$ and F are not necessarily intervals, but may be any real sub-unitary subsets. They are set-valued vector functions or operations depending on known or unknown parameters and may be continuous or discrete. Moreover, $T, I,$ and F may overlap or be converted from one to the other. An element A (t, i, f) belongs to the set in the following way: it is t true $(t\ e\ T)$, i indeterminate $(i\ e\ I)$ and f false $(f\ e\ F)$, where $t, i,$ and f are real numbers in the sets $T, I,$ and F [12].

IV. CLOUD DETECTION ALGORITHM-PROPOSED APPROACH

Remote sensing images are very important remote sensing data sources. Due to the influence of weather, cloud cover is one of the common obstacles for data analysis in the satellite images. So we propose an automatic approach for cloud detection. The proposed algorithm consists of following steps: Color transformation, Wavelet decomposition, Texture extraction, Clustering and Morphological operations.

Block Diagram of proposed method is shown below.
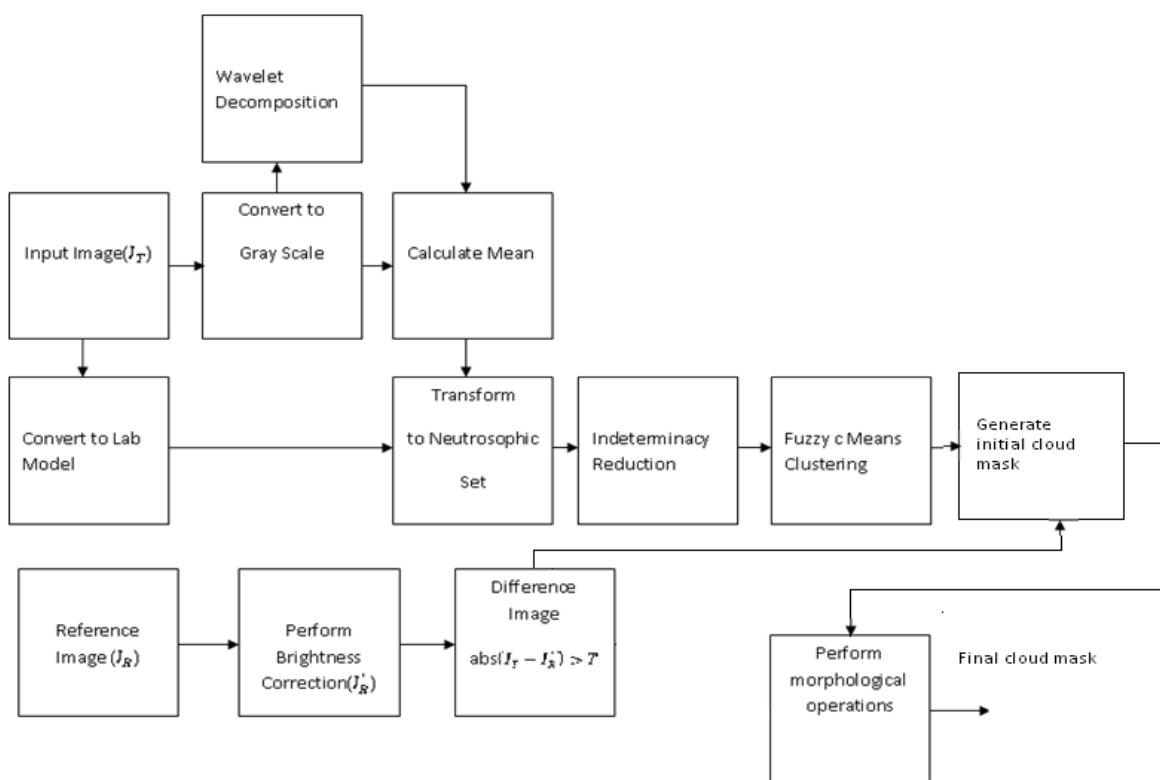
Fig 1 Block diagram

The algorithm is summarized as below:

Step 1: Convert image from RGB color space to Lab color space.

Step 2: Obtain gray scale image from RGB color space.

Step 3: Use wavelet transformation to decompose the gray scale image into sub-bands (LL, LH, HL, and HH).

Step 4: Calculate the mean of LH, HL sub-bands and the gray scale image using Eqs. (1)– (3).

Step 5: Transform the L color channel and the mean values into NS domain independently using Eqs. (4)– (8).

Step 6: Perform the indeterminacy reduction operation on the true subsets using Eqs. (9) and (10).

Step 7: Apply Fuzzy C-means clustering to the true subset using Eq (11).

Step 8: Find the absolute difference between input and brightness corrected reference image.

Step 9: Perform Morphological operations.

In this paper, we propose a cloud detection algorithm based on wavelet transform and neutrosophic sets [13,14]. The clouds are detected based on the color and texture features. Cloudy satellite images generally contain the combination of color and texture. Therefore, combining color and texture features would be of significant benefit in distinguishing regions having the same color but different textures, and vice versa. In the first step, RGB input image is converted into Lab color space. The color features are extracted from Lab transformed image. Lab color is designed to approximate human vision. It gives good results for segmentation of color pictures. It aspires to perceptual uniformity, and its L component closely matches human perception of lightness. So for further processes, only the L component is used. For extracting the texture features, the RGB input image is converted into gray scale image. Then, wavelet transform is applied to gray scale image to obtain the LL, LH, HL, and HH sub-bands. 2/2 biorthogonal wavelet decomposition is used [15], which is separable and computationally efficient. One-level wavelet decomposition is performed. Only, LH and HL sub-bands are employed for further processes because most of the texture information is in the LH and HL sub-bands.

Then mean of LH, HL sub-bands and the gray scale image in a local window are calculated. Window size of 5 is used. The mean values are calculated for texture characterization. Mean can be calculated using the following equations.

$$MELH(i,j) = \frac{1}{w \times w} \sum_{k=-w/2}^{w/2} \sum_{l=-w/2}^{w/2} LH(i+k, j+l) \quad (1)$$

$$MEHL(i,j) = \frac{1}{w \times w} \sum_{k=-w/2}^{w/2} \sum_{l=-w/2}^{w/2} HL(i+k, j+l) \quad (2)$$

$$MEGray(i,j) = \frac{1}{w \times w} \sum_{k=-w/2}^{w/2} \sum_{l=-w/2}^{w/2} Gray(i+k, j+l) \quad (3)$$

where w is the size of the sliding window. $MELH$, $MEHL$ and $MEGray$ are the mean values of LH, HL sub band and the grayscale image respectively.

Then the pixel P(i, j) in the image domain for the L channel and the calculated mean values are transformed into the neutrosophic domain [16,17], $P_{NS}(i, j) = \{T(i, j), I(i, j), F(i, j)\}$. T (i, j), I (i, j) and F (i, j) are the membership value belonging to true set, indeterminate set and false set, respectively. NS is used to describe the indeterminacy in the color texture image.

$$T(i,j) = \frac{\bar{g}(i,j) - \bar{g}_{min}}{\bar{g}_{max} - \bar{g}_{min}} \quad (4)$$

$$\bar{g}(i,j) = \frac{1}{w \times w} \sum_{m=i-w/2}^{i+w/2} \sum_{n=j-w/2}^{j+w/2} g(m,n) \quad (5)$$

$$I(i,j) = \frac{\delta(i,j) - \delta_{min}}{\delta_{max} - \delta_{min}} \quad (6)$$

$$\delta(i,j) = abs\big(g(i,j) - \bar{g}(i,j)\big) \quad (7)$$

$$F(i,j) = 1 - T(i,j) \quad (8)$$

where $\bar{g}(i,j)$ is the local mean value of the image. $\delta(i,j)$ is the absolute value of the difference between intensity g(i, j) and its local mean value $\bar{g}(i,j)$ at (i, j).

The indeterminacy of the NS image P_{NS} is decreased for getting better results. Indeterminacy reduction is performed on subset T of L channel and the mean values ($MELH, MEHL, MEGray$) using the following equations.

$$T_{mod}(i,j) = \begin{cases} T(i,j) & T(i,j) < 0.5 \\ \overline{T}(i,j) & T(i,j) \geq 0.5 \end{cases} \quad (9)$$

$$\overline{T}(i,j) = \frac{1}{w \times w} \sum_{m=i-w/2}^{i+w/2} \sum_{n=j-w/2}^{j+w/2} T(m,n) \quad (10)$$

where $T_{mod}(i,j)$ is the modified true set after the indeterminacy reduction operation and $\overline{T}(i,j)$ is the mean of the true set.

After the indeterminacy reduction operation, the membership set T becomes more distinct and has high contrast. So the clouds can be detected more effectively.

Finally a clustering is performed on the indeterminacy minimized data. Clustering is the process of dividing data elements into classes so that items in the same class are similar as possible, and items in different classes are as dissimilar as possible. Here Fuzzy C-means clustering is used. The elements in subset T for L color channel and the mean values are combined as the input for the Fuzzy C-means clustering [18]. The input of the clustering algorithm is a vector $X = [T_{L_{NS}}, T_{MELH_{NS}}, T_{MEHL_{NS}}, T_{MEGray_{NS}}]$. It is based on minimization of the following objective function:

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^m \|x_i - c_j\|^2 \quad ,1 \leq m < \infty \quad (11)$$

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 ith of d-dimensional measured data, c_j is the d-dimension center of the cluster, and $\|*\|$ is any norm expressing the similarity between any measured data and the center. Number of clusters is selected as 2. After experimental analysis we have found that cluster number 2 gives the best results. Then label every pixel in the image using the results from Fuzzy C-Means. For every object in the input, Fuzzy C-Means returns an index

corresponding to a cluster. Label every pixel in the image with its cluster index and the cluster contains cloud pixels are taken for further analysis. The cluster index for cloud pixels is found out and that cluster index is used for generating a binary cloud mask that contains cloud pixels represented as white and non cloud pixels represented as black.

Now we have a binary cloud mask that contains possible cloud pixels and we have to eliminate the non cloud pixels. It can be done by calculating the absolute difference between intensity of target and brightness corrected reference image. If it is greater than threshold T then those pixels are treated as actual cloud pixels. The threshold T is experimentally set to 25. Target image (I_T) and reference images (I_R) are two images which are observed at different times but cover the same region. In this paper we assume that the target and reference images are already registered. There is a slightly or heavily difference in the brightness of the same-location images acquired at different time due to the atmospheric effects, sun angles, and sensor looking angles. As a result, brightness correction [19] is performed to achieve consistency in the mean and standard deviation of the intensity values of target and reference images. In this process, the mean and standard deviation of reference image is linearly transformed into the mean and standard deviation of target image. The brightness correction is formulated as follows:

$$I'_R = \left(I_R - \mu_{I_R}\right) \times \frac{\sigma_{I_T}}{\sigma_{I_R}} + \mu_{I_T} \tag{12}$$

where I_R is the reference image, μ_{I_R} is mean of reference image, μ_{I_T} is mean of target image, σ_{I_T} is standard deviation of target image, σ_{I_R} is the standard deviation of reference image and I'_R is the brightness corrected reference image.

The cloud-contaminated pixels are excluded in the calculation of equation 12. The reference image (I_R) is image captured over same area as that of target image but on different date. We select image with least cloud cover as reference image.

In order to avoid small misclassifications in cloud detection, we perform morphological erosion operation with a structuring element square of size 3. To select thin clouds on the cloud boundaries, a dilation operation with structuring element size 9 is applied to the detected cloud mask. Finally a binary decision map of same size as input image is created to record the detected results of locations of cloud regions.

V. EXPERIMENTAL RESULTS

The data considered in this work used a set of images acquired with the Landsat-7 ETM+ sensor. All images used in this work are cropped to 512×512 pixels.

Landsat is one of the most widely used satellite data sources for local, regional, and global applications because of its medium spatial resolution, multispectral bands, and long record of historical data. Landsat images provide the longest satellite observations of the Earth's land surface that are important to a wide range of remote sensing applications. However, due to the nature of optical sensors and the relative

long revisit cycle. Thick-cloud contamination is a common problem in Landsat images, which limits their utilities in various land surface studies. Landsat images are highly affected by clouds, which present a serious obstacle in their applications, especially in monitoring land surface dynamics.

In figure 2 (a)-(e) the different cloudy Landsat images and corresponding cloud detected images are shown. The detected cloud regions are enclosed in red boundary. The results are shows that the proposed cloud detection algorithm can accurately detect almost all thick and thin clouds.

The evaluation of cloud masks is difficult because ground truth is not available to compare the cloud mask. As no ground truth is available, the performance of the proposed algorithm is evaluated subjectively. Performance evaluation is performed quantitatively by the numbers of clouds detected correctly (NOC) and the precision of clouds detected (POC) [7].

The precision of the clouds detected is calculated as follows:

$$POC = \frac{NOC}{NOM} * 100 \tag{13}$$

where NOM is the number of clouds detected manually.

In fig 3 (a)-(b) the different cloudy Landsat images (i) and corresponding cloud detected images by proposed algorithm (ii), direct thresholding (iii),and by Wang et.al (iv) are shown. The detected cloud regions are enclosed in red boundary in (ii) and (iv) and in magenta boundary in (iii). To detect clouds by thresholding, threshold value is experimentally set to 180 and only red band of Landsat image is used. The results are shows that the proposed cloud detection algorithm can accurately detect almost all thick and thin clouds.

Table I record the total number of clouds detected by the proposed algorithm (NOC), total number of clouds detected manually (NOM) and precision (POC) of clouds detected.

TABLE I.

EVALUATION RESULTS OF CLOUD DETECTED RESULTS OF FIG 3(a)-(b)

Image	NOC	NOM	POC (in %)
a	3	3	100
b	6	6	100

The images used in this result analysis are of size 256×256 pixels. In fig 3(a) all clouds are detected, number of clouds detected manually is 3 and number of clouds detected by the algorithm is also 3 and result in 100 % precision.

Table II represents the total number of pixels detected as cloud and total number of regions detected as cloud by the proposed method, thresholding and Wang et.al mehod. In fig 3a (ii) total number of pixels detected as cloud by proposed approach is 1296, in fig 3a (iii) clouds detected by thresholding is 354 and in fig 3a (iv) cloud pixels detected by Wang et.al's method is 554. Only a small number of pixels are detected as cloud by thresholding and by Wang et.al method.
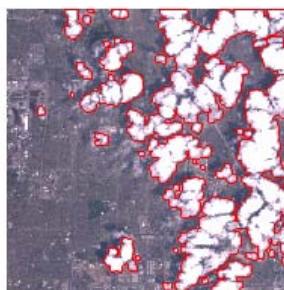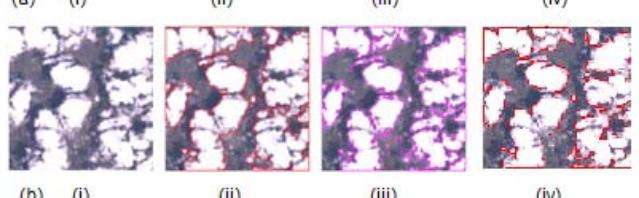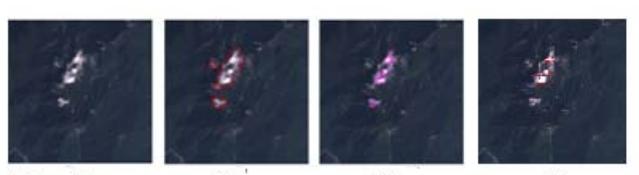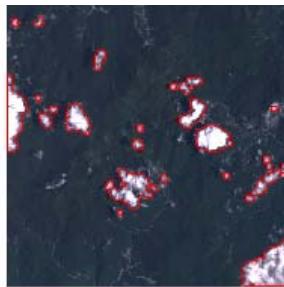
Fig 3(a) (b)(i) Cloudy input image corresponding (ii) cloud detected image by proposed approach (iii) thresholding (iv) and by Wang et.al

TABLE II.

EVALUATION RESULTS OF CLOUD DETECTED RESULTS OF FIG 3(a)-(b)

Image	Total Number of pixels detected as Cloud			Total Number of regions detected as Cloud		
	Proposed Approach	Thres holding	Wang et.al	Proposed Approach	Thres holding	Wang et.al
a	1296	354	554	5	8	3
b	30657	32916	40073	27	62	14

In fig 3b (ii) total number of pixels detected as cloud by proposed approach is 30657, in fig 3b (iii) clouds detected by thresholding is 32916 and in fig 3b (iv) clouds detected by Wang et.al is 40073. Here in fig 3b (iii) and (iv) some non cloud pixels are also detected as cloud. In fig 3a (ii) total number of regions detected as cloud by proposed approach is 5, in fig 3a (iii) clouds detected by thresholding is 8 and by Wang et.al is 3. Thresholding results in cloud segments contain only few pixels. So the number of regions detected is more than that of the proposed method.

VI. CONCLUSION

This paper presents a new approach for detecting clouds in Landsat satellite imagery. The approach uses neutrosophic set and wavelet transformation. Then mean of the sub bands and grayscale image are calculated for extracting the texture features. The experimental results shows that our proposed approach yield good detection result, subjectively and quantitatively. The proposed algorithm is tested with Landsat images. The experimental results have demonstrated that the proposed algorithm classifies cloud pixels more accurately in time.

ACKNOWLEDGMENT

Original Images   (e)   Cloud Detected Images

Fig 2 (a) (b) (c) (d) (e) Cloudy input images and corresponding cloud detected images

REFERENCES

[1] J. Ju and D. P. Roy, "The availability of cloud-free Landsat ETM Plus data over the conterminous United States and globally," Remote Sens. Environ., vol. 112, no. 3, pp. 1196–1211, 2008.

[2] B. Wang, A. Ono, K. Maramatsu, N. Fujiwara, "Automated detection and removal of clouds and their shadows from Landsat TM images", IEICE Trans. Inform. Syst. E82-D (2), 453–460, 1999.

[3] Azimi-Sadjadi, M.R. ,Shaikh, M.A. ; Bin Tian ; Eis, K.E. ; Reinke, D.,'Neural network-based cloud detection/classification using textural and spectral features ',Geoscience and Remote Sensing Symposium,. IGARSS. 'Remote Sensing for a Sustainable Future.', International Page(s): 1105 - 1107 vol.2 1996.

[4] Salvador, Gabarda. S. and Cristobal. G., "Cloud covering denoising through image fusion. Image and Vision Computing", Vol . 25, pp. 523–530, 2007.

[5] Irish, R.R., "Landsat 7 automatic cloud cover assessment Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery", In Proceedings of SPIE: Orlando, FL, USA VI, 24–26, pp. 348–355 , 2000.

[6] Jiao. Q., Luo. W., Liu. X. and Zhang. B., "Information reconstruction in the cloud removing area based on multitemporal. Remote Sensing and GIS Data Processing and Applications", SPIE Vol. 6790, 679029,2007

[7] Yu Tong, WANG Ming-Shu, "Clouds and Shadows Detection in Multi-Spectral Satellite Image Based on Maximally Stable Extremal Regions", International Conference on Multimedia and Signal Processing, 2011.

[8] C. Huang, N. Thomas, S. N. Goward, J. G. Masek, Z. Zhu, J. R. G. Townshend, and J. E. Vogelmann, "Automated masking of cloud and cloud shadow for forest change analysis using Landsat images," Int. J. Remote Sens., vol. 31, no. 20, pp. 5449–5464, 2010.

[9] Bin Tian ,Azimi-Sadjadi, M.R. ; Haar, T.H.V. ; Reinke, D.,"Neural network-based cloud classification on satellite imagery using textural features ",Image Processing, 1997. Proceedings., International Conference on Page(s): 209 - 212 vol.3.

[10] Y. Guo, H.D. Cheng, "A New Neutrosophic Approach to Image Segmentation", Pattern Recognition, vol. 42, pp. 587-595, Oct. 2009.

[11] F. Smarandache, "A Unifying Field in Logics: Neutrosophic Logic. Neutrosophy, Neutrosophic set, Neutrosophic Probability", 3rd ed., American Research Press, 2003.

[12] H. Wang, R. Sunderraman, F. Smarandache, Y.Q. Zhang, "Interval neutrosophic sets and logic: theory and applications in computing", Infinite Study, 2005.

[13] Abdulkadir Sengur, Yanhui Guo., "Color texture image segmentation based on neutrosophic set and wavelet transformation," Computer Vision and Image Understanding, vol. 115, pp. 1134-1144, April 2011.

[14] M. Khoshnevisan, S. Singh, "Neurofuzzy and neutrosophic approach to compute the rate of change in new economies, in: F. Smarandache (Ed.)", Proceedings of the First International Conference on Neutrosophy, Neutrosophic Logic, Neutrosophic Set, Neutrosophic Probability and Statistics, University of New Mexico, pp. 56–62, 2002.

[15] A. Cohen, I. Daubechies, J.C. Feauveau, "Biorthogonal bases of compactly supported wavelets", Communications on Pure and Applied Mathematics, vol.45, pp. 485–560, 1992.

[16] H.D. Cheng, Y. Guo, "A new neutrosophic approach to image thresholding", New Mathematics and Natural Computation, vol. 4, pp. 291–308, 2008.

[17] Y. Guo, H.D. Cheng, "A new neutrosophic approach to image denoising", New Mathematics and Natural Computation vol.5, pp.653–662, 2009.

[18] James C. Bezdek,Robert Ehrlich,William Full,"FCM: The Fuzzy C-Means Clustering Algorithm",Computers & Geosciences Vol. 10, No. 2-3, pp. 191-203, 1984.

[19] Caselles, V., and Lopez Garcia, M.J. (1989), An alternative simple approach to estimate atmospheric correction in multitemporal studies. Int. J. Remote Sens. Vol 10, pp:1127-1134, 1989.