iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 3 **Issue:** VII **Month of publication:** July 2015

DOI:

www.ijraset.com
IC Value: 13.98
Volume 3 Issue VII, July 2015
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Efficient Segmentation of Remote Sensing Images Using New Clustering Algorithm

Reshma.M[1], P. Lakshmi Devi[2] M.Tech.,(P.hd)
Dept of ECE, AITS College of Engineering and Technology Rajampeta.

Abstract: Segmentation of real-world remote sensing images is challenging because of the large size of those data, particularly for very high resolution imagery. For segmentation of remote sensing images, many algorithms have been proposed, to provide accurate results of segmentation by using this new proposed model. Here segmentation can be done by using improved 2D gradient histogram and MMAD (minimum mean absolute deviation) model. This proposed algorithm comes under 'Thresholding ' , the optimal threshold value can find by using MMAD model. Experiments on remote sensing images indicate that the new algorithm provides accurate segmentation results, particularly for images characterized by Laplace distribution histograms.

I. INTRODUCTION

A. Segmentation of image

Image segmentation refers to the partitioning of an image into non-overlapping different regions with similar attributes. The basic attribute of gray level image is luminance amplitude, for color or multispectral images, color or information components are used. There are lot of methods are invented for segmentation like
Edge based methods (Zuraj & Lattuti,1998)
Region growing methods (Tremeau & Bonel,1998)
Neural network methods
Physics based methods (Maxwell & shafer, 1996) and
Histogram cluster thresholding methods(Sezgin & sankur, 2004).
The Histogram cluster thresholding method is good candidate for achieving segmentation for wide class of gray level images with low computational complexity. For color or multi spectral images gradient value is used to extract the information. The image gradient is directional change in intensity or color in an image. Our proposed algorithm we use the improved 2D-gradient histogram.

B. Previous Algorithms

Numbers of clustering algorithms are proposed for segmentation of remote sensing images. From these, best clustering algorithms are

1) Otsu algorithm: The Otsu scheme, a widely used image thresholding technique, provides approving results for segmenting a gray level image with only one modal distribution in gray level histogram. However it provides poor results if the histogram of a gray level is non-bimodal. Suppose the intensity of a gray level image be expressed in L gray levels[1,2,……,L]. The number of points can be expressed as $X = x_1+x_2+\ldots\ldots.+x_L$. The histogram of this gray level image is regarded as a occurrence distribution of probability

$$P(i) = \frac{xi}{X},\ x_i \geq 0,\ \sum_{i=1}^{L} xi = 1 \qquad (1)$$

The image pixels are divided into two parts C_0 and C_1, i.e. foreground and background by a threshold t. Where C_0 represents pixel within levels [1,2…..t], and C_1 denotes pixels within levels[t+1,…..,L]. The occurrence probabilities of this class and average can be expressed as respectively.

$$\omega_0 = \omega(t) = \sum_{i=1}^{t} p(i) \qquad (2)$$

$$\omega_1 = 1-\omega(t) = \sum_{i=t+1}^{L} p(i). \qquad (3)$$

$\mu_0 = \sum_{i=1}^{t} \frac{i.p(i)}{\varphi 0} = \frac{1}{\varphi(t)} \sum_{i=1}^{t} i.p(i)$ (4)

$\mu_1 = \sum_{i=t+1}^{L} \frac{i.p(i)}{\varphi 1} = \frac{1}{1-\varphi(t)} \sum_{t+1}^{L} i.p(i)$ (5)

Total mean can be written as

$\mu_T = \sum_{i=1}^{L} p(i)$ (6)

And we can find that

$\mu_T = \varphi_0\mu_0+\varphi_1\mu_1$ (7)

where φ_0 and φ_1 denote probabilities of foreground part and background part. Besides μ_0 , μ_1 and μ_T refer to the mean in gray levels of the foreground of the gray image, the background of the gray image, and the enter gray level image.

The between-class variance σ_B^2 of the two classes C_0 and C_1 is given by

$\sigma_B^2 = \varphi_0(\mu_0-\mu_T)^2 +\varphi_1(\mu_1-\mu_T)^2$ (8)

The separable degree ŋ of the class, in the discrimination analysis, is

$Ŋ = \max_{1<t<L} \sigma_B^2$ (9)

Finally, maximizing σ_B^2 choose the optimal threshold t^*

$t^* = \arg_{1<t<L}\max\sigma_B^2$ (10).

2) Minimum within class variance algorithm (MWCV): Otsu suggested minimizing the weighted sum of within class variances of the foreground and background pixels to establish an optimum threshold. Recall the minimization of within-class variance is tantamount to the minimization between-class scatter. This method given satisfactory results when the numbers of pixels in the each class are close to each other. The Otsu method still remains one of the isodata clustering is given in Velasco. Some limitations of the Otsu method are disclosed in Lee and Park. Liu and Li generalized it to a 2D Otsu algorithm.

3) K-means clustering algorithm: K-means algorithm is an unsupervised clustering algorithm that classifies the input data points into multiple classes based on their inherent distance from each other.
The algorithm assumes that the data features from a vector space and their to find natural clustering in them, the points are clustered around centroids μ_i for all i=1,2,….,k which are obtained by minimizing the objective.

$V = \sum_{i=1}^{K} \sum_{xj \in si} (xj - \mu i)^2$ (1)

Where there are K clusters S_i, i=1,2….,K and μ_i is the centriod or mean point of all the points $x_i€ S_i$.
As a part of this project an iterative version of the algorithm was implemented. The algorithm takes a two dimensional image as input various steps in the algorithm as follows

a) Complete the intensity distribution (also called the histogram) of the intensities.
b) Initialize the centroids with K-random intensities.
c) Repeat the following steps until the cluster a label of the image does not change any more.
d) Cluster the points based on distance of their intensities from the centroids intensities.

$C^{(i)} = \arg\min_j //X\, i - \mu j//^2$ (2)

e) the new centroid for each of clusters

Volume 3 Issue VII, July 2015
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

$$\mu_i := \sum_{i=1}^{m} K\{(Ci = j)\}Xi \,/\, \sum_{i=1}^{m} K\{Ci = j\} \quad (3)$$

Where K is a parameter of the algorithm (the number of clusters to be found) i iterates over the all the intensities, j iterates over all the centroids and μ_i are the centroids intensities.

4) Minimum Error Thresholding algorithm: These methods assure that the image can be characterized by a mixture distribution of foreground and background pixels. P(g) = P(T).
$P_f(g)+[1-P(T)]P_b(g)$ Lloyd considers equal variance Gaussian density functions and minimizes the total mis-classification error via an iterative search. In contrast Kittler and Illingworth removes the equal variance assumption and, in essence, addresses a minimum error Gaussian density fitting problem. Recently Cho, Haralick and Yi have suggested an improvement of this thresholding method by observing that the means and variances estimated from truncated distribution result in a bias. This bias becomes noticeable however, only whenever the two histogram modes are distinguishable.

a) Drawbacks in above algorithms: When applying these algorithms into remote sensing image, there are certain drawbacks described in the following
- *i)* In the 2D histogram there are too many points having low values in the remote sensing images; therefore, computing all points is a significiant wast of computational time.
- *ii)* Treditional histograms only take grayscale information into consideration when introducing this type of histogram into remote sensing images the objects are segmented incompletely because remote sensing images contain complex texture information.
- *iii)* Certain proposed algorithms perform well on images characterized by a standard Gaussian distribution, such as Otsu and MET. However, when images characterized by other distribution histograms are the mixture of Gaussian distribution are not standard the segmentation results are not satisfactory.

II. PROPOSED ALGORITHM

In this letter we propose a novel thresholding algorithm to segment the roads and residential areas from vegetation areas in remote sensing images. In consideration of both flatness and grayscale information, we construct an improved 2D gradient histogram. To decrease the computational complexity, global features are extracted from the 2D histogram as a 1D histogram by diagonal projection. The MMAD model can obtain an accurate optimal threshold based on global feature. Experimental results indicate that this algorithm performs well for remote sensing images, not only for those characterized by Gaussian distribution histograms but also for those characterized by laplacian distribution histogram. In addition, the time consumption of our algorithm is accepted.

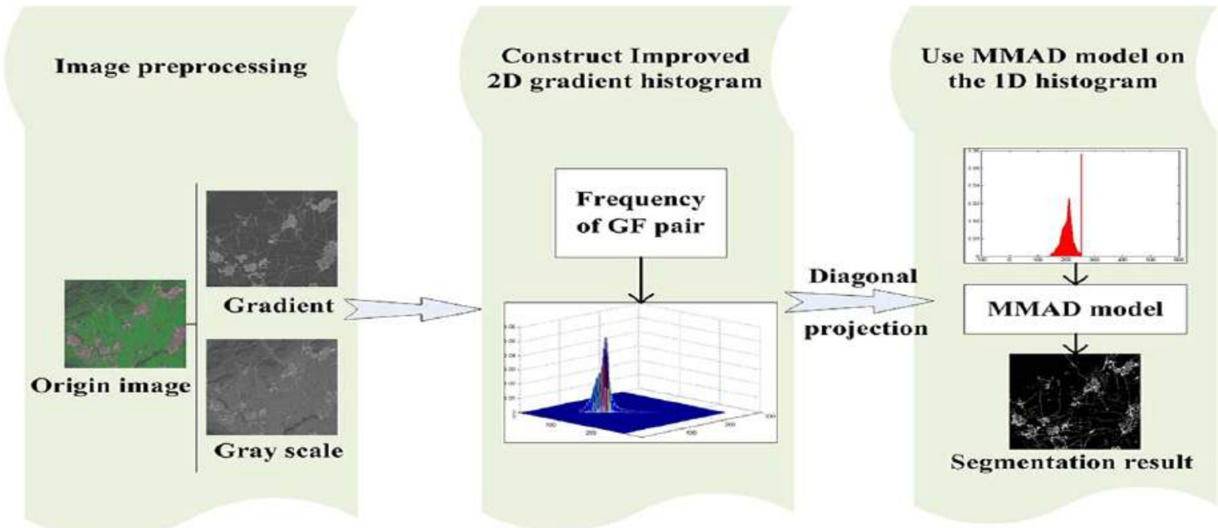


Fig: Frame work of proposed algorithm

The input image is processed to obtain the gradient image and the gray scale image for constructing an improved 2D gradient histogram. Next, the global featured are extracted as a 1D histogram from 2D histogram by diagonal projection. Framework can be divided into three sections

A. Improved 2D gradient Histogram

Let I denote a grayscale image with L gray levels of size M×N. The 2D histogram is composed of two parameters of the image; one is grayscale image denoted by f(x,y) and the gradient of f at the location of (x,y) is given by

$$\Delta f = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right]^{T} \quad (1)$$

For discrete image, gradient can be generated from the running differences of pixels along vertical and horizontal axes of the image i.e., given as

$$\frac{\partial f}{\partial x} = f(x+1, y) - f(x-1, y) \quad (2)$$

$$\frac{\partial f}{\partial y} = f(x, y+1) - f(x, y-1) \quad (3)$$

Consider both horizontal and vertical gradients equally, the united gradient g(x,y) indicates as follows:

$$g(x, y) = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} \quad (4)$$

Each pixel in the images has its united gradient(G) and grayscale(F) as its GF pair.
Let P_{ij} be the frequency of GF pair (i,j) where g(i,j) = i and f(i , j) = j, then

$$P_{ij} = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \delta ij \quad (5)$$

Where

$$\delta ij = \begin{cases} 1, g(x, y) = i \text{ and } f(x, y) = j \\ 0, otherwise \end{cases}$$

B. Transformation from the 2D to 1D histogram

The below figure(a) shows the GF pair of remote sensing images are mainly concentrated on the diagonal of 2D histograms. The 2D histogram matrix is given by

$$P_r = \sum_{i+j=r} Pij \quad (6)$$

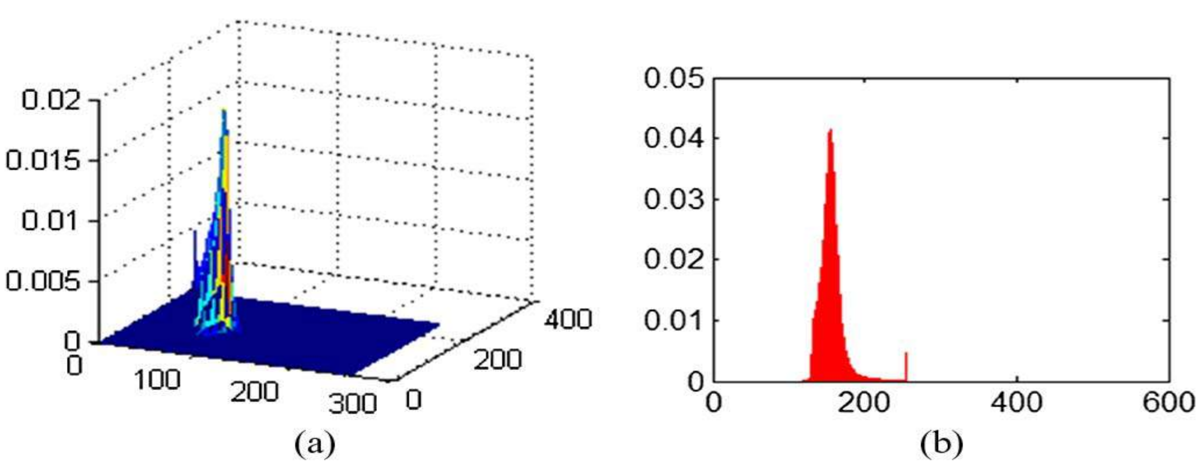


Fig:(a) projection of Improved 2D gradient histogram ,and (b) projection of 1D histogram.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Where i is abscissa in the improved 2D gradient histogram and j is the ordinate in the improved 2D gradient histogram. Hence the 2D parameter space is reduced to a 1D histogram of the variable r, where r$\in \{0,1,2,\dots\dots,2(L-1)\}$.

C. MMAD Model

In the 1D Histogram, let k be an assumed threshold for binary segmentation. For binary segmentation, pixels in the images are divided into two classes: one class for pixels corresponding to bins r r$\in (0,1,2..,k)$ in the 1D histogram, and the other class corresponding to bins r$\in \{k+1,k+2,\dots,2(L-1)\}$ in the 1D histogram. For a given threshold k, the class probability and class mean are as follows:

$$\begin{cases} \omega o(k) = \sum_{r=0}^{k} Pr \\ \omega 1(k) = \sum_{r=k+1}^{2(L-1)} Pr \end{cases}$$

$$\begin{cases} \mu 0(k) = \sum_{r=0}^{k} rp(r) \\ \mu 1(k) = \sum_{r=k+1}^{2(L-1)} rp(r) \end{cases}$$

Next, we can construct the mean absolute deviation on 1D histogram i.e.,

$$\begin{cases} MAD0(k) = \sum_{r=0}^{k} \frac{|r-\mu o(k)|pr}{\omega o(k)} \\ MAD1(k) = \sum_{r=k+1}^{2(L-1)} \frac{|r-\mu 1(k)|pr}{\omega 1(k)} \end{cases}$$

The proposed algorithm we selects the threshold by minimizing

$$\text{MAD(k)} = \sum_{i=0}^{1} MADi(k)$$

11

i.e., k* = MMAD= $\arg_{k\in\{0,1,2,\dots,L-1\}} \min[MAD(k)]$

Where k* is the optimum threshold value is used for effective segmentation.

D. Theoretical interpretation

Our algorithm performs well for remote sensing images, and it is more suitable when the histograms correspond to the Laplace distribution. We provide the theoretical interpretation to this phenomenon.

The probability density function(pdf) of the Laplace distribution is as follows

$$f\left(\frac{x}{\mu},\sigma\right) = \frac{1}{\sqrt{2\sigma}} e^{-\frac{|x-\mu|}{\sigma\sqrt{2}}}$$

Where μ is the sample mean, and σ is the sample variance.

In the 1D histogram, the two scenes each can appear as a bell-shaped mode, and the threshold can be selected as the valley between modes. The modes of two scenes are denoted by $f(x/\mu 0,\sigma 0)$ and $f(x/\mu 1,\sigma 1)$, respectively. Each probability of $f(x/\mu 0,\sigma 0)$ and $f(x/\mu 1,\sigma 1)$ is $\omega 0(x)$ and $\omega 1(x)$, and the function of the 1D shown as follows:

$$h(x) = \omega 0(x)\, f(x/\mu 0,\sigma 0) + \omega 1(x)\, f(x/\mu 1,\sigma 1)$$

where $0\le x\le 2(L-1)$. Hence, for the optimal threshold k*, k* satisfies

$$\omega 0(k^*)\ f(k^*/\mu 0,\sigma 0) = \omega 1(k) f(k^*/\mu 1,\sigma 1).$$

From h(x) the optimal threshold k* is

Volume 3 Issue VII, July 2015
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

$$K^* = \frac{\sigma 0 \sigma 1/\sqrt{2}\ln\left(\frac{\psi 0 \sigma 1}{\psi 1 \sigma 0}\right)+\mu 0 \sigma 1+\mu 1 \sigma 0}{\sigma 0+\sigma 1}.$$

In the remote sensing image the prominent difference of the gray scale and the gradient between the two types of scenes, there is peak with a very small variance at the high value in the 1D histogram i.e., $\sigma 1 << \sigma 0$ then from above k* equation $k^* \rightarrow \mu 1$. For histograms of mixture of Laplace distribution $f(x/\mu o, \sigma o)$ and
$f(x/\mu 1, \sigma 1)$ where $\sigma 1 << \sigma 0$.
Then the mean absolute deviation is as follows:

$MAD(k)=MAD_0(k)+MAD_1(k)$

Where

$$MAD_0(k) = \sum_{x=0}^{k}\left[f\left(\frac{x}{\mu 0},\sigma 0\right)+f\left(\frac{x}{\mu 1},\sigma 1\right)\right]|x-\mu 0(k)|,$$

$$MAD_1(k) = \sum_{x=0}^{k}\left[f\left(\frac{x}{\mu 0},\sigma 0\right)+f\left(\frac{x}{\mu 1},\sigma 1\right)\right]|x-\mu 1(k)|$$

and small value of $\sigma 1$ $f(x/\mu 1,\sigma 1)$ is very small i.e., near to zero ,except at $x=\mu 1$.

PDF of the Laplace distribution $f\left(\frac{x}{\mu},\sigma\right), \int_{-\infty}^{\infty} f\left(\frac{x}{\mu},\sigma\right)|x-a|dx$ is minimum at $\mu=a$. Then the above MAD(k) becomes

$$\sum_{x=0}^{k} f\left(\frac{x}{\mu 0},\sigma 0\right)|x-\mu 0(k)| + \sum_{x=k+1}^{2(L-1)} f\left(\frac{x}{\mu 0},\sigma 0\right)|x-\mu 1(k)|$$

From this we can obtain the minimum optimal threshold value minimized for good results.

III. EXPERIMENTAL RESULTS

To evaluate the performance of the proposed algorithm, several experiments were conducted using selected remote sensing images. The experimental results were designed to compare the overall performances between different clustering algorithms.

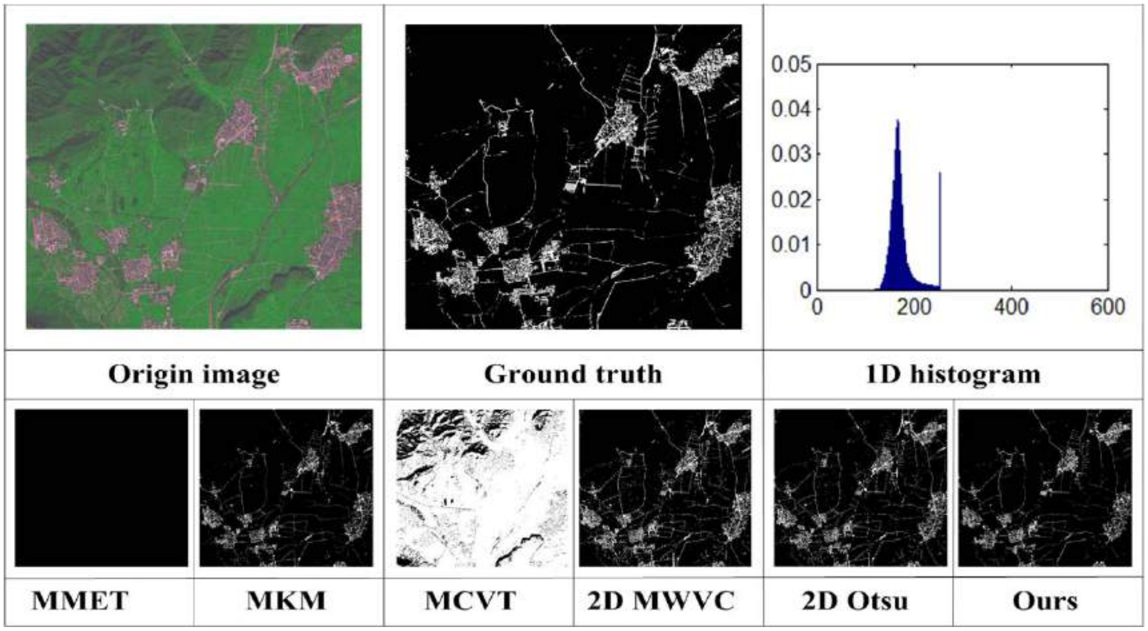


Fig:Segmentation results of various algorithms.

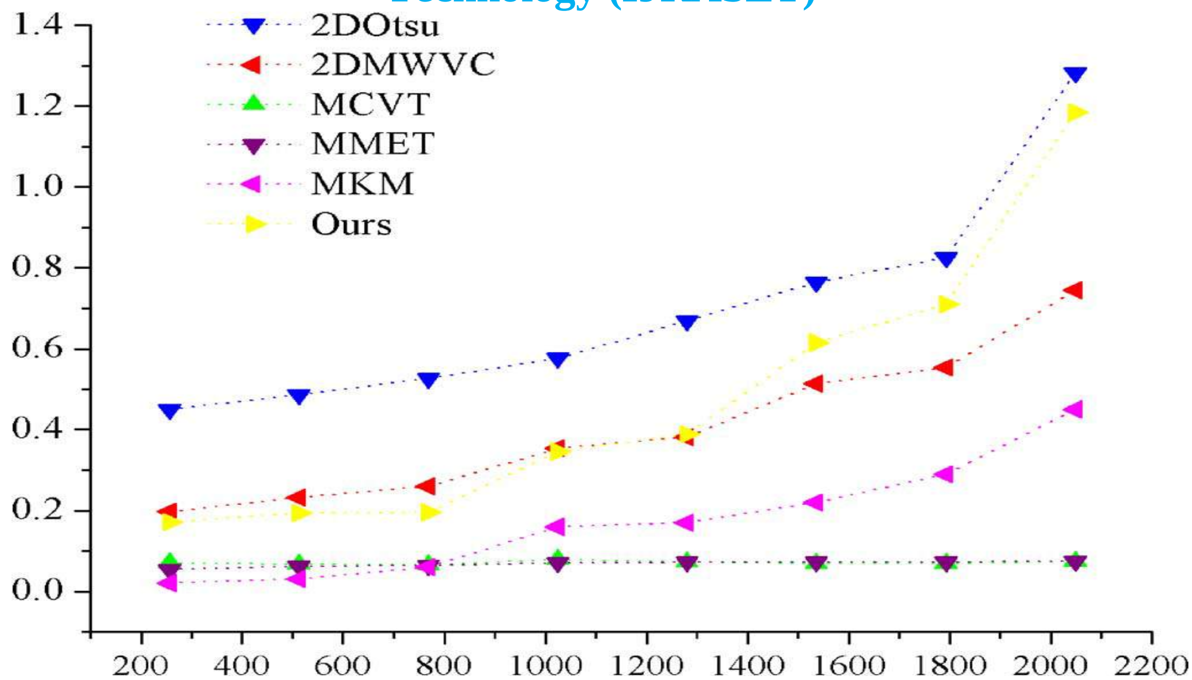


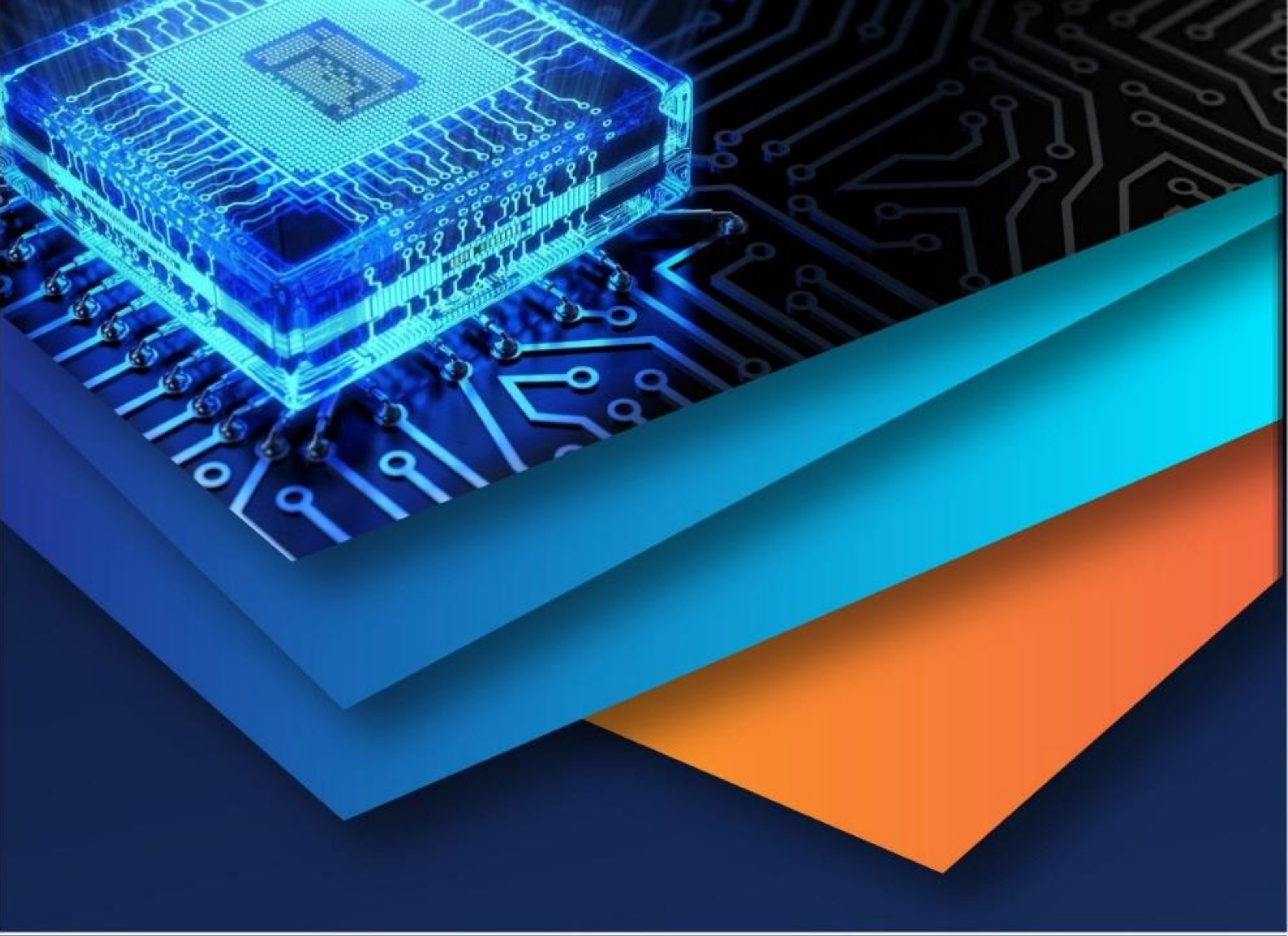
Fig: Time trends for computing the optimal threshold of various algorithms.

IV. CONCLUSION

In this paper we proposed and validated algorithm for segmentation of remote sensing images. The features of the input image is extracted from an improved 2D gradient histogram as a 1D histogram. The new clustering model i.e., MMAD model is used on to obtain the optimal threshold. The segmentation results of proposed algorithm are statistically satisfactory. This algorithm not only expands thresholding segmentation of remote sensing images characterized by Laplace distribution histogram but also meets the time requirement.

REFERENCES

[1] N. Otsu, "A threshold selection algorithm from gray-level histograms,"IEEE Trans. Syst. Man Cybern., vol. SMC-9, no. 1, pp. 62–66, Jan. 1979.
[2] M. Sezgin, "Survey over image thresholding techniques and quantitative performance evaluation," J. Electron. Imaging, vol. 13, no. 1, pp. 146–165, Jan. 2004.
[3] J. Xue and D.M. Titterington, "Median-based image thresholding," Image Vis. Comput., vol. 29, no. 9, pp. 631–637, Aug. 2011.
[4] Q. Chen et al., "Modified two-dimensional Otsu image segmentation algorithm and fast realization," IET Image Process., vol. 6, no. 4, pp. 426–433, Jun. 2012.
[5] F. Nie, J. Li, T. Tu, and P. Zhang, "Image segmentation using two dimensional extension of minimum within-class variance criterion," Int. J. Signal Process., Image Process. Pattern Recognit., vol. 6, no. 5, pp. 13–24, Oct. 2013.
[6] S. Cho, R. Haralick, and S. Yi, ''Improvement of Kittler and Illingworths's minimum error thresholding,'' Pattern Recogn. 22, 609–617~1989!.
[7] J. Kittler and J. Illingworth, ''On threshold selection using clustering criteria,'' IEEE Trans. Syst. Man Cybern. SMC-15, 652–655 ~1985!.
[8] N. Otsu, ''A threshold selection method from gray level histograms,'' IEEE Trans. Syst. Man Cybern. SMC-9, 62–66 ~1979!.
[9] C. V. Jawahar, P. K. Biswas, and A. K. Ray, ''Investigations on fuzzy thresholding based on fuzzy clustering,'' Pattern Recogn. 30~10!,1605–1613 ~1997!.
[10] F. R. D. Velasco, ''Thresholding using the isodata clustering algorithm,''IEEE Trans. Syst. Man Cybern. SMC-10, 771–774 ~1980!.
[11] H. Lee and R. H. Park, ''Comments on an optimal threshold scheme for image segmentation,'' IEEE Trans. Syst. Man Cybern. SMC-20,741–742 ~1990!.
[12] J. Z. Liu and W. Q. Li, ''The automatic thresholding of gray-level. pictures via two-dimensional Otsu method,'' Acta Automatica Sin. 19, 101–105 ~1993!.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY