

Image Histogram Segmentation by Multi-Level Thresholding using Hill Climbing Algorithm

Sayantana Nath
Research Scholar
Indian Institute of Information
Technology, Allahabad, India

Dr. Sonali Agarwal
Assistance Professor
Indian Institute of Information
Technology, Allahabad, India

Qasima Abbas Kazmi
M.Tech Student
Indian Institute of Information
Technology, Allahabad, India

ABSTRACT

Image Segmentation is an essential technique in image processing area to distinguish important object from unnecessary background substrates. Most of the image segmentation methods are based on the “Cloud Histogram” or Density Variation Concept which cannot be capable to work with individual value of the histogram of image. The “Hill Climbing” based Multilevel Thresholding technique will overcome the limitation and it is applicable to the value of image histogram directly to recognize the absolute pitch point turn over. This technique is based on individual value of histogram column. The highest folds of the image histogram curve play a major role in graphical resolution and visual orientation. This research perspective is not given importance in the field of histogram clustering yet.

Keywords

Image Segmentation, Hill Climbing Algorithm, Differential Tangent Equation, Local Minima, Local Maxima.

1. INTRODUCTION

Image thresholding is very useful tool for separating figures from the background substrate with discriminate object contains different color level. The variety of thresholding techniques, among global histogram based algorithms is employed to determine the thresholds told by a survey. In parametric approach of gray level distribution for each class has a probability density function which obeys the Gaussian distribution attributes. An attempt to find out the estimated parameters over the distribution will be fit the given histogram data by using the least square method best. Especially it is followed by a non-linear optimization problem whose solution is computationally infeasible and time consuming due to exponential values of standard deviation and mean variance of the Gaussian density distribution function. In the view of image segmentation by multilevel thresholding there exist no such algorithms which can work with the histogram value directly. The basic technique to find out the multilevel threshold values is the Gaussian density distribution function. But the limitation of this technique is bounded only in the field of “Cloud Histogram Concept” and it cannot be operated with the individual each value of the image histogram columns. In an alternative algorithm fitting curve based on heuristics method is used to detect tree annealing [1]. Genetic algorithm based First Scheme Approach is implemented for optimal thresholding [2]. But the Multivariate Automated Thresholding Criterion [3] is one of the robust multilevel thresholding techniques in image segmentation area. Recently a hybrid and heuristics Hill climbing technique

[6] mutated with the both Nelder-Mead simplex search algorithm [4] and particles swarm optimization abbreviated method as (NM – PSO) [5] is proposed to solve the objective function of Gaussian fitting curve for multilevel thresholding. But these entire algorithms are strictly based on the Cloud histogram techniques. There are some constraints to calculate the thresholds of image histogram processing. In the previous techniques the lowest occurrence of the color level pixel is taken as threshold point of histogram [7]. The lowest folds of the histogram curvature are considered as the set point for image clustering [8]

2. HILL CLIMBING AND HISTOGRAM HYBRIDIZATION

2.1 Mean Value Calculation

1st threshold value, T1 : - the highest no. of occurrence of pixel image.

2nd threshold value, T2: the least no. of occurrence of the pixel image.

No. of divided region, $N \leq$ (position of the highest frequency of the pixel, $p_1 \sim$ position of the least frequency of pixel, p_2) / constant value, C.

$$\text{Where constant value, } C = \left| \frac{(v_2 p_1 - v_1 p_2)}{p_1 - p_2} \right|.$$

Where, v_1 = the occurrence value of the pixel at the p_1 .

v_2 = the occurrence value of the pixel at the p_2 .

$$\text{The range of the each divided region is, } R = \frac{(p_2 \sim p_1)}{N}.$$

The position of the each divided sector of the histogram by the N is = C_t where $t \in I^+$ upto N-1.

Where $C_0 = p_1$ and $C_N = p_2$. So, $C_{s+1} - C_s = R$, where $s \in I^+$ upto N.

The coordinate of the $C_x = p_1 \pm x * R$ at the any position of the xth divided partition where $x \in t$. So, $C_N = p_1 \pm N * R = p_2$ or, $p_2 \sim p_1 = N * R$.

The median_position_value, $M_w = (the\ 1st\ coordinate\ position\ taken\ C_w + the\ 2nd\ coordinate\ position\ taken\ C_{w+1}) / 2$;

2.2 Local Minima and Maxima Calculation

So,

$$M_w = [(p_1 \pm w * R) + \{p_2 \pm (w - 1) * R\}] / 2$$

where, $w \in I^+$ upto N.

Or, $M_w = [C_w + C_{w+1}] / 2$

Then, $M = \cup M_d$, where $d \in I^+$ upto N.

In the Hth divided section by the Hill Climbing Algorithm we get:

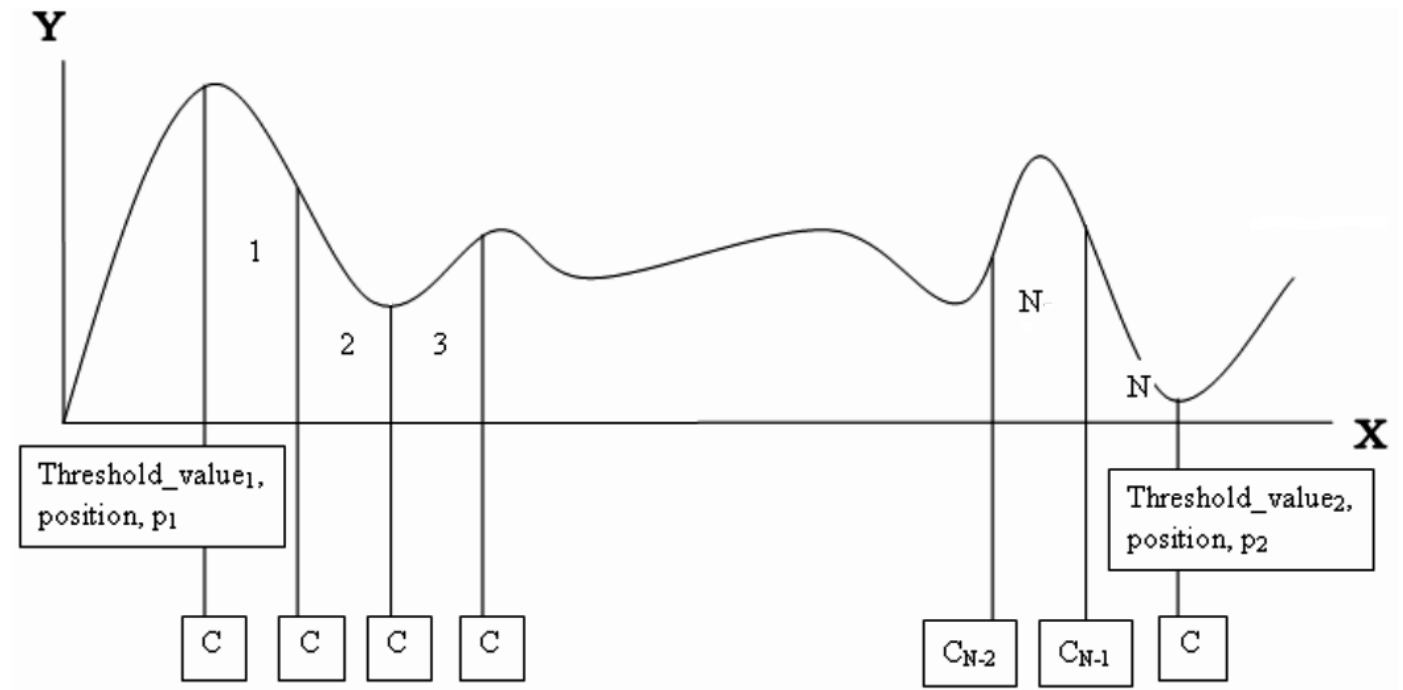


Fig 1: Total number of intermediate C calculation.

The above histogram is divided into the N no. of section in where $C_1, C_2 \dots C_{N-1}$ is denoted as the coordinate of the limit

The local minima, l_{h_1} = between C_{h-1} and M_h for the p_1 at that side

The local maxima, l_{h_2} = between C_h and M_h for the p_2 at that side.

Where, $h \in I^+$ upto N

Then the union of:

$$\left. \begin{array}{l} \text{The local minima, } L_1 = \cup M_{k_1} \\ \text{The local maxima, } L_2 = \cup M_{k_2} \end{array} \right\} \begin{array}{l} \text{Where, } k \in I^+ \\ \text{upto } N \end{array}$$

The calculation of value of C shown in the **Figure 1** below:

of each range. The respective value is shown in the **Figure 2** below:

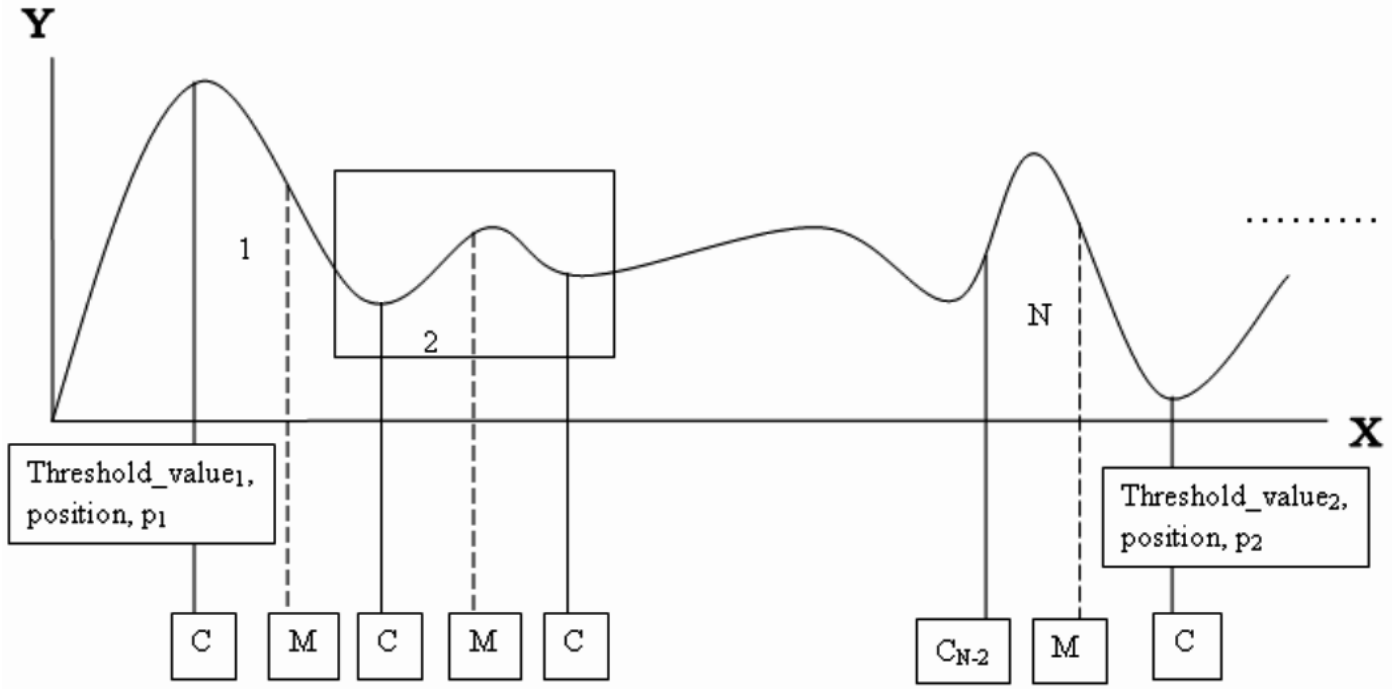


Fig 2: Smallest derivational value in-between the position of pair of Cs

The above histogram is shown the median points in between section of the each divided segment. Now, the box region on the

before image is shown below in the magnified view. The respective result is shown in the **Figure 3** below:

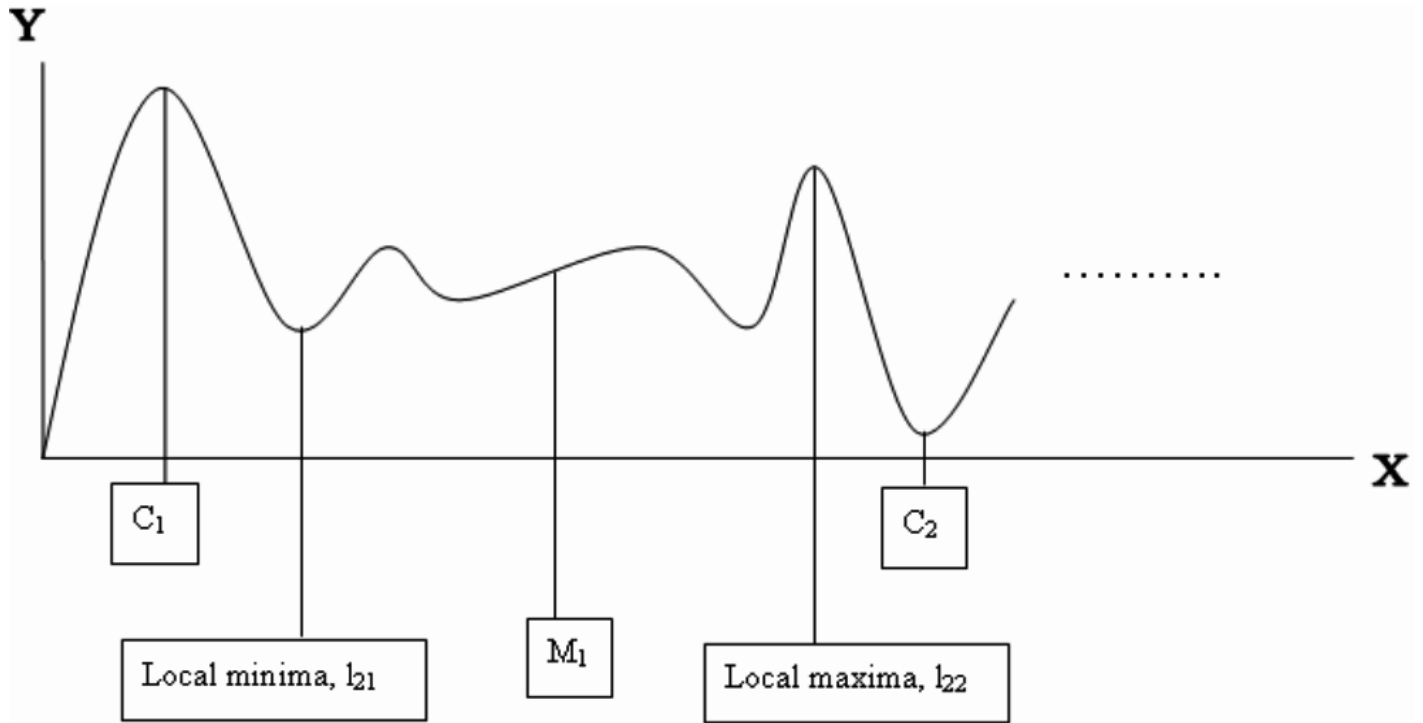


Fig 3: Capital Median calculation between highest and lowest pixel histogram

Now, if the no. of division between p_1 and p_2 is, $N \rightarrow \infty$. So the in between range of the each pair of the divided coordinate position is, $C_{x+1} - C_x$ or $R \rightarrow 0$ where $x \in I^+$ upto N and

$C_{x+1} - C_x$ are denoted as dx . So, $dx \rightarrow 0$. The manipulated result is shown in the **Figure 4** below:

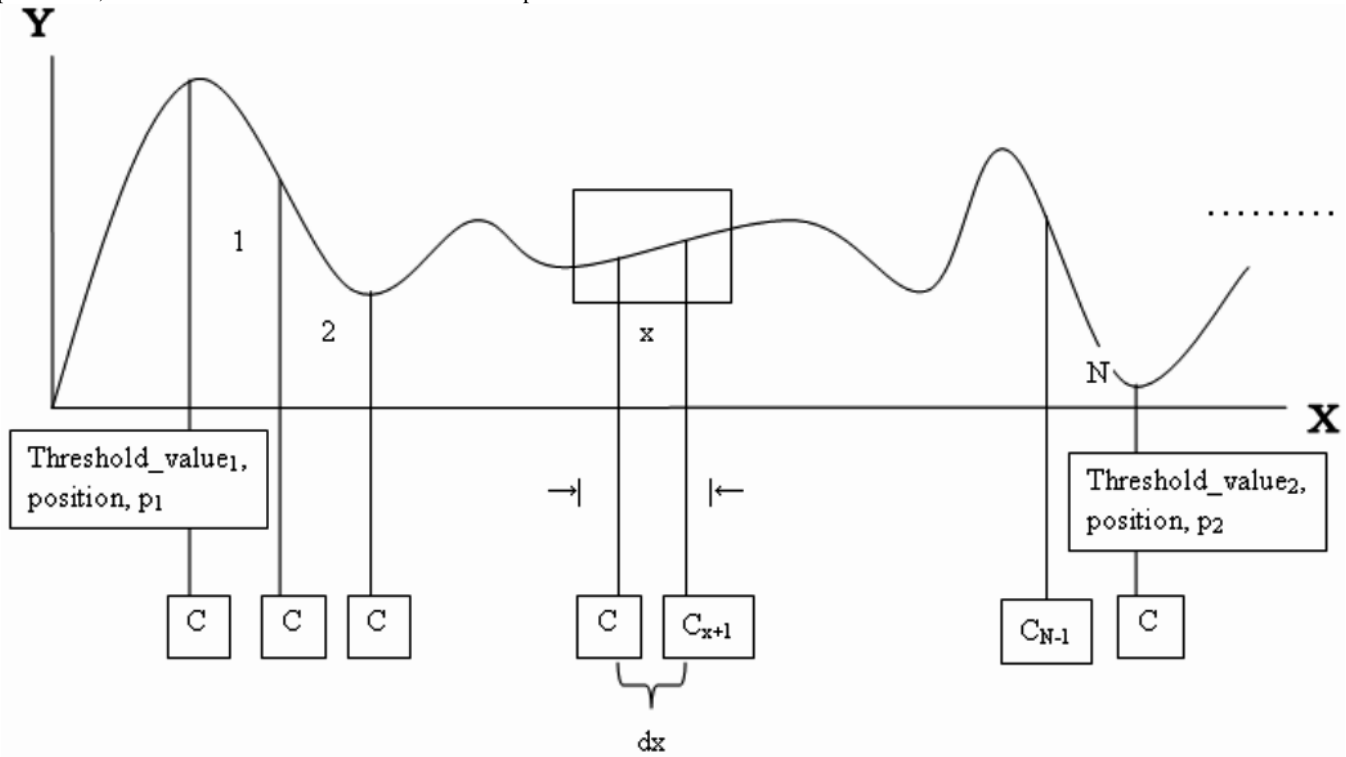


Fig 4: variation of mean calculation between every pair of C

Now, the enlarge section of the square box shown in the **Figure 5** below:

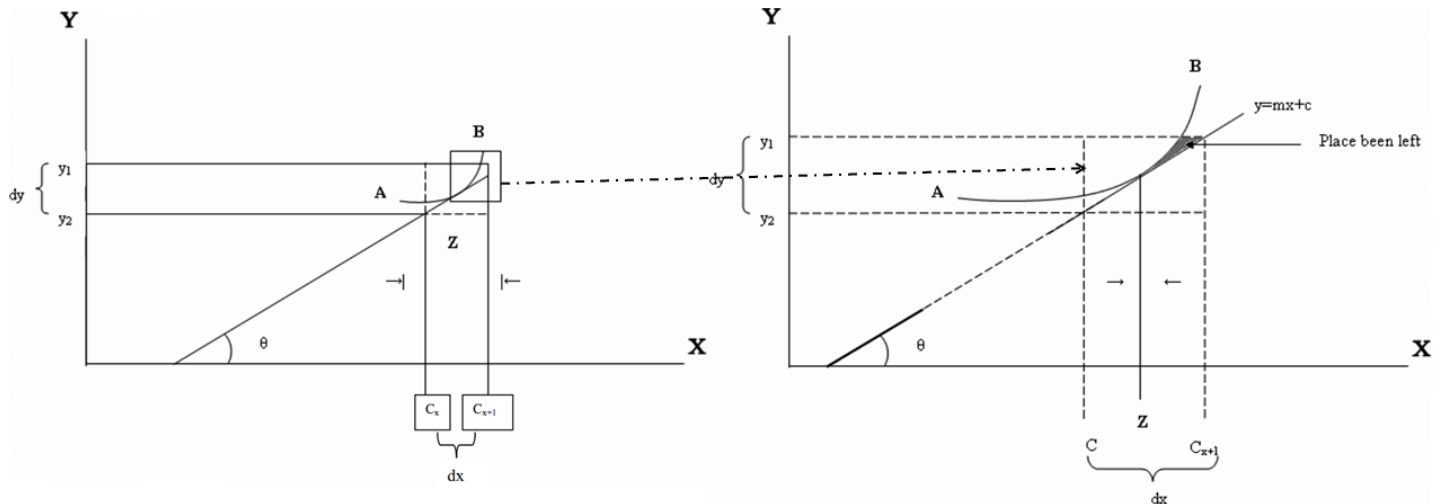


Fig 5: Tangential Angle measurement with respect to X – axis to calculate the hyperbolic considerable value of the histogram curvature

2.3 Comparison between Integral and Hill Climbing Constrains

The AB curve portion of the full histogram shown in the before figure has a tangent straight line through the point Z creating an angle θ^0 with respect to X axis.

So, the equation of the straight line is, $y = mx + c$ where $m = \tan \theta \dots (1)$

And we know $\tan \theta = \frac{dy}{dx} \dots (2)$

Then, comparing (1) and (2) we get: -

$$m = \frac{dy}{dx} .$$

Now, putting the value in the equation we get: -

$$y = \left(\frac{dy}{dx}\right)x + c;$$

$$\text{or, } y \, dx = x \, dy + c \, dx;$$

$$\text{or, } x \, dy = (y - c)dx.$$

$$\text{or, } \frac{dy}{y-c} = \frac{dx}{x};$$

$$\text{or, } \int \frac{dy}{y-c} = \int \frac{dx}{x};$$

The range of the y axis is v_1 to v_2 and the range of the x axis is p_1 to p_2 .

$$\int_{v_1}^{v_2} \frac{dy}{y-c} = \int_{p_1}^{p_2} \frac{dx}{x}$$

$$\text{Or, } \log(y - c) \Big|_{v_1}^{v_2} = \log(x) \Big|_{p_1}^{p_2}$$

$$\text{Or, } \log(v_2 - c) - \log(v_1 - c) = \log(p_2) - \log(p_1)$$

$$\text{Or, } \log \frac{v_2 - c}{v_1 - c} = \log \frac{p_2}{p_1}$$

$$\text{Or, } \frac{v_2 - c}{v_1 - c} = \frac{p_2}{p_1}$$

$$\text{So, } c = \left| \frac{v_2 * p_1 - v_1 * p_2}{p_1 - p_2} \right|$$

Where, $c \equiv C$.

So, in the section of $dx * dy$ region the tangent $y = mx + c$ is passing through the point Z of the curvature AB. Now at the situation of the no. of division $N \rightarrow \infty$ so, $(c_x - c_{x+1}) \rightarrow 0$ or, $dx \rightarrow 0$, but the division points do not merge with the perpendicular from the point Z on the x axis since if it merges then $dx = 0$ so, $\tan \theta = \infty$ which can not be solved. So, there exists some gap or place left in the area of the $dy * dx$ in the each segmented region which are divided. For that the total

summation of the left areas are the calculation point for the farther proceedings of the iteration and required segmented parts which are denoted by the extra constant value calculated from the straight line equation of the tangent, c. this is the proceeding value of the extra as same as the left portion area of the histogram in each segment. This extra constant value is to be calculated for the margin of the necessary no. of segmentation of the histogram in the total range in between the p_1 and p_2 . So, the calculated region must be between the v_1 to v_2 for y axis and x_1 and x_2 for the x axis.

2.4 Proposed Scheme

Threshold_value_selection ()

{ Threshold_value1=the highest no. of occurrence of the pixel in the image; Threshold_value2=the least no. of occurrence of the pixel in the image;

Where v_1 =the occurrence value of the pixel at the p_1 . v_2 =the occurrence value of the pixel at the p_2 .

No._of_divided_region, $N \leq (\text{position_highest_frequency_pixel}, p_1 \sim \text{position_least_frequency_pixel}, p_2) / \text{constant_value}, C$.

Where constant value, $C = \left| \frac{(v_2 * p_1 - v_1 * p_2)}{(p_1 - p_2)} \right|$.

The range of the each divided region is, $R = \frac{(p_2 - p_1)}{N}$.

For(each_section)

{ Median_position_between(C_x, C_{x+1}), $M_x = [\text{any_selected_division_position}, p_1 \pm x * R + \text{next_selected_division_position}, p_1 \pm (x+1) * R] / 2$;

Local_minima_at_highest_side, $m_1 = \text{between}(\text{median_position}, p_1 \pm x * R)$; Local_maxima_at_least_side, $m_2 = \text{between}(\text{median_position}, p_1 \pm (x+1) * R)$; }

2.5 Approach of Hill Climbing Algorithm

2.5.1 Finding of Local minima at the highest occurrence value

Between_range (median_position_between(C_x, C_{x+1}) or M_x, C_x)

{ Local_counter₁= Local_counter₂=0;

Comparison_position₁=median_value_position, M_x ;

For (i=median_value_position to C_x)

{ If [$f(i-1) < f(\text{Comparison_position}_1)$]

{ Comparison_position₁= Comparison_position₁ - 1;

Local_counter₁<=0;

Local_counter₂<=0;

}

```

If [f (i-1) = f (Comparison _position1)]
Local_counter1++;
Else
Local_counter2++;
If (Local_counter1=> C OR Local_counter2=> C)
Exit (loop);
}
Threshold_value3=f (comparison_position1);
}

```

2.5.2 Finding of Local maxima at the least occurrence value

Between_range (median_position_between(C_x , C_{x+1}) or M_x , C_{x+1})

```

{
Local_counter1= Local_counter2=0;
Comparison_position2=median_value_position, Mx ;
For (j=median_value_position to Cx+1)
{
If [f (j+1) > f (Comparison_position2)]
{
Comparison_position2= Comparison_position2 + 1;
Local_counter1<=0;
Local_counter2<=0;
}
If [f (j+1) = f (Comparison_position2)]

```

```

Local_counter1++;
Else
Local_counter2++;
If (Local_counter1=> C OR Local_counter2=> C)
Exit (loop);
}
Threshold_value4=f (comparison_position2);
}
C=constant value defined before.

```

Where, f(x) = value of the no. of occurrence of the pixel at the 'x' color level position.

And, constant_value= mean value between threshold_point1 and the median, $\left\lfloor \frac{(p_2-p_1)}{4} \right\rfloor$.

3. EXPERIMENTAL RESULT

Typically, it leads to a non-linear optimization problem, the Hill Climbing Algorithm the initial value of calculation used from the median value between the least and highest no. pixel frequency at their cardinality. The Hill Climbing Algorithm is applied to image histogram and the operation of the local minima and maxima concept to find out the median of the threshold value. The proposed algorithm is employed on a remote sensing image of fire in a coal seam at the District of Jhariya, [9] Jharkhand, India. This Hill Climbing Based Histogram Segmentation technique helps to find out the respective area of fire spreading. The fire affected region is manipulated to calculate the fire intensity. The temperature increasing rate can be also recognized in this process. The original image is shown in the **Figure 6** below:



Fig 6: Coal Seam Fire at Jhariya District in Jharkhand

The following histogram shows the red color orientation only.

The histogram is shown in the **Figure 7** below:

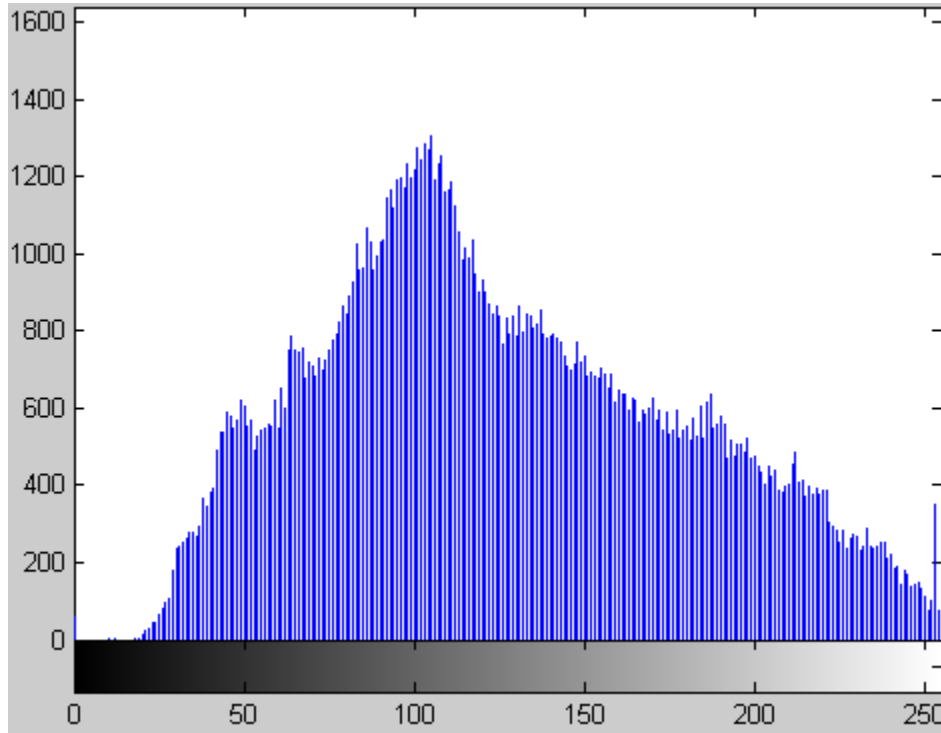


Fig 7: Red color level histogram

The following histogram shows the Green color orientation only.

The histogram is shown in the **Figure 7** below:

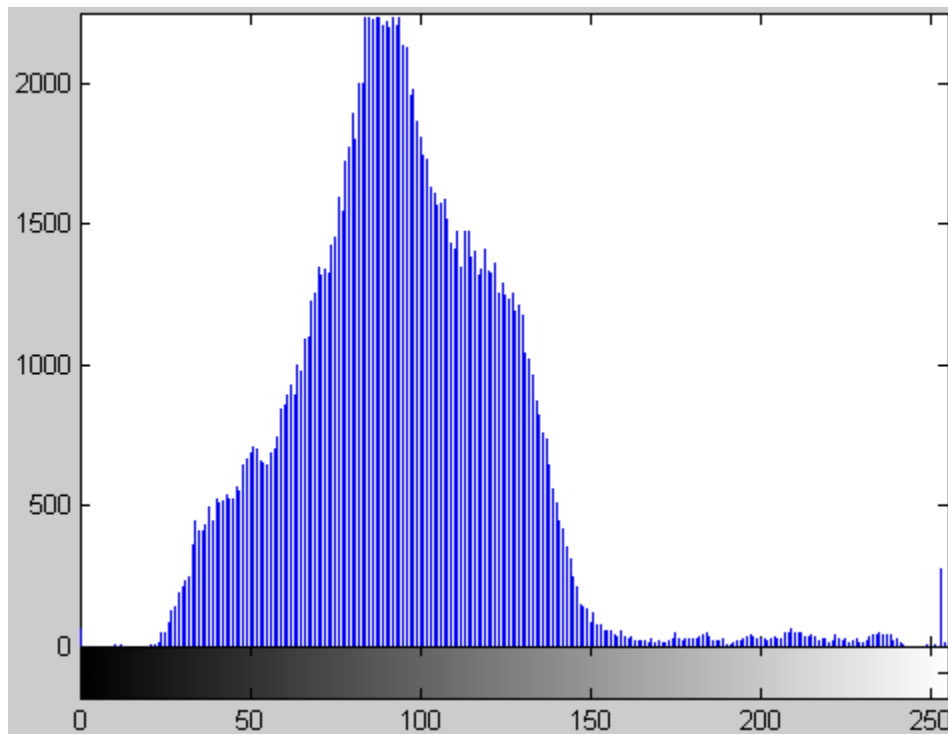


Fig 7: Green color level histogram

The following histogram shows the Blue color orientation only.

The histogram is shown in the **Figure 8** below:

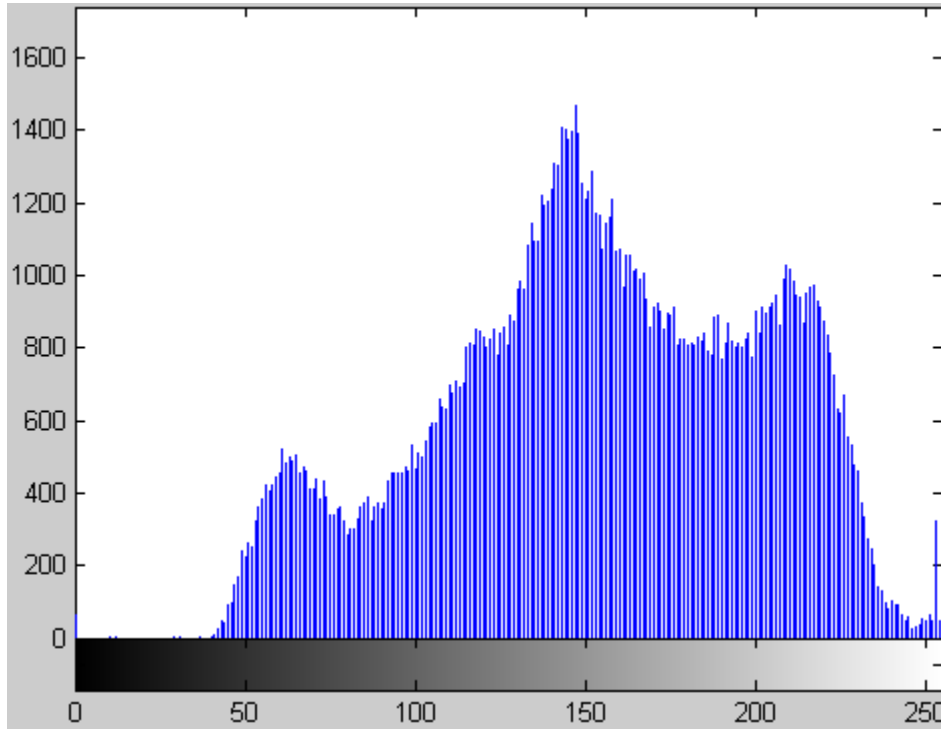


Fig 7: Blue color level histogram

The coal fire image, taken by remote sensing, is segmented by the proposed hill climbing algorithm based multilevel histogram thresholding respectively as blue, green and red.

The blue color based segmentation shows the heat affected area due to coal seam fire. The respective area shown in the **Figure 7** below:

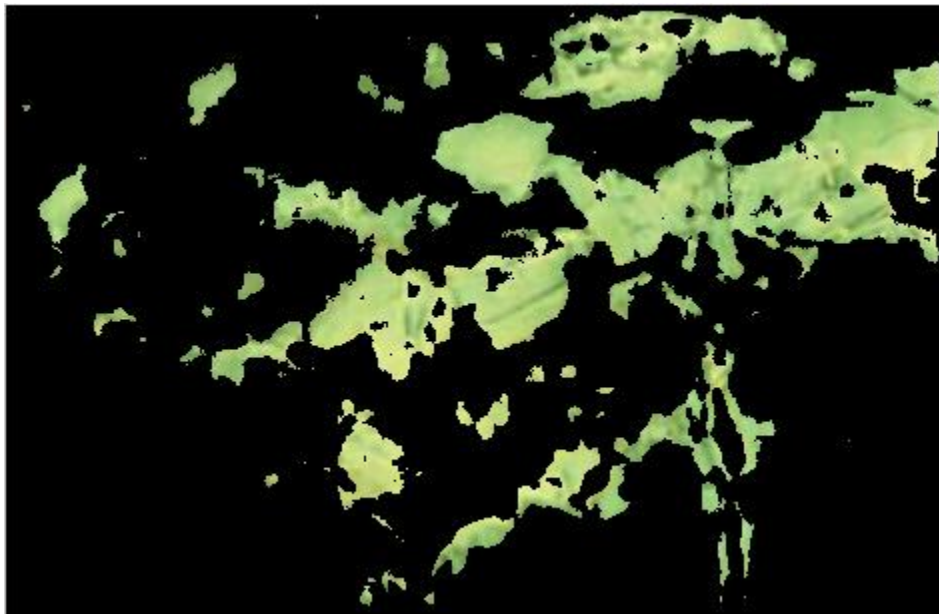


Fig 7: Respective Area of Temperature Increasing

The green color based segmentation represents the probable fire spreading region at the mine area.

The respective area shown in the **Figure 8** below:

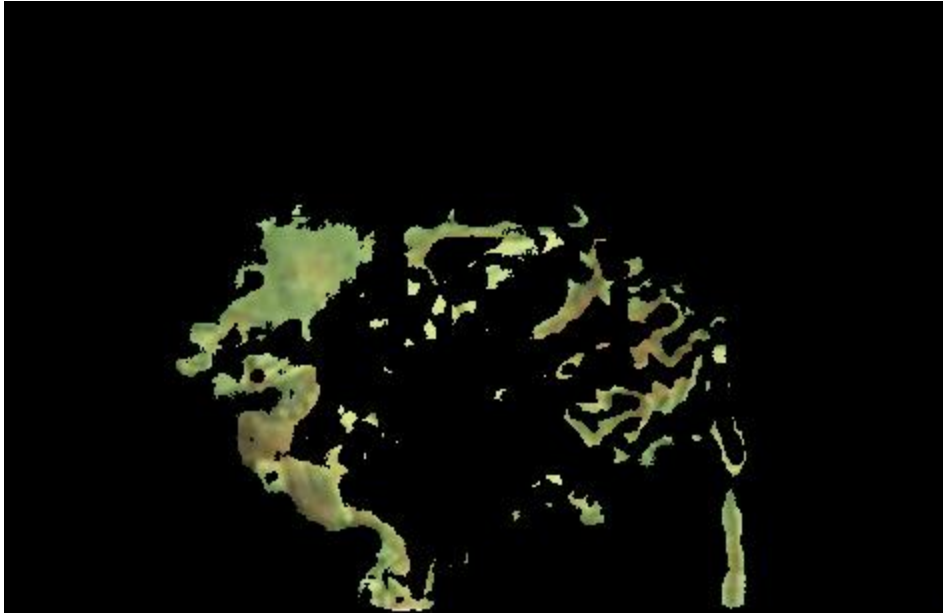


Fig 8: Possible Fire Spreading Area

The red color based segmentation shows direct fire area at the coal seam of Jharkhand .

The respective area shown in the **Figure 9** below:

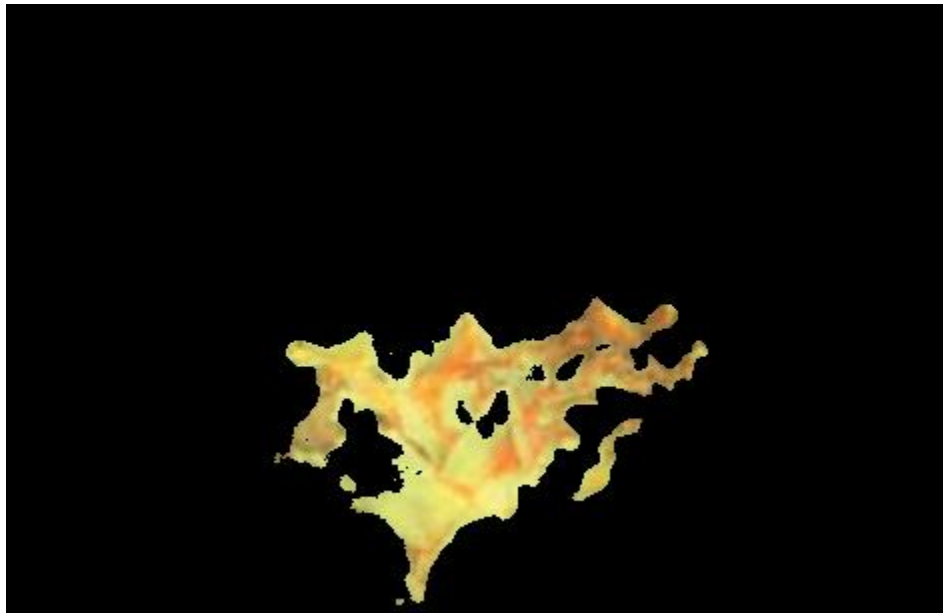


Fig 8: burning fire area at Jhariya coal field

An attempt to find out estimated parameters of the distribution will fit the best in given histogram data by using least-square method.

4. CONCLUSION

The Gaussian Density Distribution function or the Gaussian Fitting Curve method is very effective in the thresholding case, but its computational time becomes deviated in case of multilevel thresholding and it does not work with the individual values of the histogram. To make the evolutionary computation method in the field of independent image segmentation with more perspective, we have proposed a more visible thresholding point selecting scheme called the “Multilevel Thresholding with Hill Climbing Algorithm” that solves the multilevel segmentation process better than Gaussian curve fitting and Gaussian density distribution function in term of computational time complexity. The application of the proposed scheme is implicated on the simulator, MATLAB. To sum up, the Hill Climbing Algorithm with the local minima-maxima concept is a promising and effective tool for image segmentation by multilevel histogram thresholding due to its computation efficiency, and this method proves the effectiveness with respect to the quality performance than previous schemes.

5. ACKNOWLEDGMENTS

Our thanks is oriented to Krishna Kant Agarwal, Vimal Upadhyay and Akhilesh Srivastav who have contributed towards development of the template.

6. REFERENCES

- [1] Glasbey, Yin Zuink, C.A., “An Analysis of Histogram Based Thresholding Algorithm”, 2003.
- [2] Fan.S.K, Tsai Yen, S.K.Gonsal, “A New Criterion for Automatic Multilevel Thresholding”, 2008.
- [3] Zaraha E, F.J.Chang and G.Bilbro, “Optimal Multi-Thresholding Using Hybrid Optimization Approach”, 2005.
- [4] Renders J.M, S.P.Flasse, “Hybrid Method for Genetic Algorithm for Global Optimization”, 2006.
- [5] Alexandre Temporel, Tim Kovacs; “A heuristic hill climbing algorithm for Mastermind”, <http://www.cs.bris.ac.uk/Publications/Papers/2000067.pdf>
- [6] Kishor Bhoyar, Omprakash Kakde; “Color Image Segmentation Based On JND Color Histogram”; *International Journal of Image Processing (IJIP)* Volume(3), Issue(6).
- [7] Stéphanie Jehan-Besson, Michel Barlaud, Gilles Aubert, Olivier Faugeras; “Shape Gradients for Histogram Segmentation using Active Contours”; *Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003)*, Volume – 2.
- [8] Dr. Sonali Agarwal, and Prof. G.N. Pandey, “Studies on E Governance in India using Data Mining Perspective”, *Journal of Computing*, Volume 2, Issue 10, October 2010
- [9] Source of Application image – [“http://bio.ltsn.ac.uk/imagebank/default.aspx”](http://bio.ltsn.ac.uk/imagebank/default.aspx)