AN EXTENSIVE SURVEY OF IMAGE SEGMENTATION TECHNIQUES

Sarvesh Karandikar¹, Aditi Singh², D. Malathi³, J.D. Dorathi Jayaseeli⁴

¹ B.Tech Student, Department of Computer Science and Engineering, SRM University
² B.Tech Student, Department of Computer Science and Engineering, SRM University
³ Professor, Department of Computer Science and Engineering, SRM University
⁴ Assistant Professor, Department of Computer Science and Engineering, SRM University

ABSTRACT
The paper describes in detail the techniques used for edge detection in images as well as evolutionary algorithms which can be applied for Multi-level thresholding. In the first section it describes the core idea of edge detection followed by the first derivative operators being used currently for edge detection like: Roberts, Sobel, Prewitt, Robinson and Kirsch. The next section continues with explanation regarding the second derivative operators like: Laplacian, of Gaussian, Difference of Gaussians and Marr-Hildreth Algorithm. Further, it is followed by techniques used to detect edges using region based method known as Multi-Level Thresholding which is being applied using either the Otsu or the Kapur method. Evolutionary Algorithms namely Particle Swarm optimisation, Firefly Optimisation Algorithm and Elephant Herding Algorithm are explained in context with Kapur’s method for Multi-Level Thresholding. Finally, experimental results regarding one common image and application of the first and second derivative edge detection techniques is observed in relation with calculation of Mean Square Error (MSE), Root Mean Square Error (RMSE) and Peak Signal to Noise Ratio (PSNR) of the original image to the detected image.

Keywords: Edge Detection, First Derivative, Kapur, Multi-level Thresholding, Second derivative

I. INTRODUCTION
Image segmentation is a process which partitions input image into set of multiple segments such as a set of pixels so that the resulting image is easier to analyze [1]. It is usually used to locate boundaries, edges or regions having same characteristics. It is a process by which pixels having same properties such as texture, color, and intensity belong to one region. Different regions have different characteristics. Image segmentation algorithms can be categorized into two different approaches: discontinuity and similarity. Discontinuity based approach partition’s the image based on abrupt changes in intensity. This method includes point, line or Edge detection. Similarity based approach partition’s an image based on regions that are similar according to some predefined criteria. This approach includes Thresholding, Region growing, Region splitting and merging. In our paper we have studied in depth about edge detection methods, and different operators used in it and optimization techniques to find optimal solutions for Multi-Level Thresholding problem.
II. EDGE DETECTION

Edge detection aims at identifying points of discontinuities in digital images or points where intensity changes abruptly. The points at which the brightness changes sharply are organized into a set of curved line segments termed edges. Edge is defined as a boundary between two homogeneous regions with distinct grey level properties. This process detects the outline of an object and boundaries between objects [2]. We are keen to detect edges because most of the information pertaining to an image is enclosed within the edges. Edge detection helps in simplifying the analysis of the given image and reduces amount of data to be processed. It also helps in preserving useful information such as object boundaries of the image. So to process an image first we detect images by using appropriate filters and then by enhancing those areas of images containing edges we increase its sharpness and make it clearer. Edge detection technique has various applications such as biometrics, face detection, satellite image segmentation etc. It is useful in research field because it helps in analyzing images more efficiently [3]. Its techniques are classified into two types: first derivative technique and second derivative technique.

III. FIRST DERIVATIVE OPERATORS

First order derivative operators work on detecting the start and the end of an edge based on varying intensities of pixels present in the grayscale image. If there is drastic change in intensity between two pixels the operators mark the given pixel as an edge. First derivative operators tend to produce thicker edges. They are based on the value being obtained from the 2D gradient after application of the algorithm. The gradient is calculated at a specific point x,y and is generally obtained with a vector associated at the particular point. The most commonly used discontinuity based edge detection techniques are reviewed in this section. Those techniques are Roberts edge detection, Sobel Edge Detection, Prewitt edge detection, Kirsch edge detection, Robinson edge detection.[4]

A. Roberts Edge Detection

Roberts edge detection method performs the operation of calculation of gradient on the image in the X and the Y direction of the image associated. In this method the regions with high frequency are the ones which generally constitute that of an edge in the image. The input as well as the output is a grayscale image which is pretty common. In the output image the pixel values generally constitute the gradient value at that given point. [4]

\[ \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \]

B. Prewitt Edge Detection

This is a mask which is used to detect horizontal as well as vertical type of edges. In this mask operator being used we have two masks available at our disposal. One is the mask which is used for horizontal direction oriented edges and the other is for the vertically embedded edges in the image.[4]

\[ \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \]
C. Sobel Edge Detection

This operator is almost same as the prewitt operator in many characteristics as it is also classified as the derivative mask. It also is used for Vertical and Horizontal edge detection but is different than prewitt operator in the sense that it does not have any fixed coefficients of the mask. The mask can be changed according to our needs and greater negative values being imparted into the mask will lead to better enhancement of the edge as compared to the prewitt operator image.[5]

X gradient
\[
\begin{bmatrix}
1 & 2 & 1 \\
0 & 0 & 0 \\
-1 & -2 & -1
\end{bmatrix}
\]

Y gradient
\[
\begin{bmatrix}
1 & 0 & -1 \\
2 & 0 & -2 \\
1 & 0 & -1
\end{bmatrix}
\]

D. Robinson Edge Detection

This operator is a collection of masks by rotating a given mask in 8 different directions of the compass namely East West North South North-East North-West South-East And South-West keeping the column /row vector of the mask zero. This implies that the edge being detected is emphasized when the detection process gets over with. As presented in the above two operators the edge detection of Robinson Compass would help carry out edge detection task pertaining to direction of edges more efficiently.

North
\[
\begin{bmatrix}
1 & 0 & 1 \\
2 & 0 & 2 \\
-1 & 0 & 1
\end{bmatrix}
\]

North west
\[
\begin{bmatrix}
0 & 1 & 2 \\
-1 & 0 & 1 \\
-2 & -1 & 0
\end{bmatrix}
\]

West
\[
\begin{bmatrix}
1 & 2 & 1 \\
0 & 0 & 0 \\
-1 & -2 & -1
\end{bmatrix}
\]

South west
\[
\begin{bmatrix}
2 & 1 & 0 \\
1 & 0 & -1 \\
0 & -1 & -2
\end{bmatrix}
\]

South
\[
\begin{bmatrix}
1 & 0 & -1 \\
2 & 0 & -2 \\
1 & 0 & -1
\end{bmatrix}
\]
E.Kirsch Edge Detection

This operator again resembles to the same characteristics as the Robinson compass mask operator but the only overall difference is the value of the coefficients present in the mask are not constant and can be changed according to our need. This freedom of applying masks of our own values of weights lets us determine the edge more efficiently and predominantly than Robinson operator. [6]

\[
\begin{bmatrix}
-3 & -3 & 5 \\
-3 & 0 & 5 \\
-3 & -3 & 5 \\
\end{bmatrix}
\]

North west

\[
\begin{bmatrix}
-3 & 5 & 5 \\
-3 & 0 & 5 \\
-3 & -3 & -3 \\
\end{bmatrix}
\]

North

\[
\begin{bmatrix}
5 & 5 & 5 \\
-3 & -3 & -3 \\
\end{bmatrix}
\]

West

\[
\begin{bmatrix}
5 & 0 & -3 \\
-3 & -3 & -3 \\
\end{bmatrix}
\]

South west

\[
\begin{bmatrix}
5 & -3 & -3 \\
-3 & -3 & -3 \\
\end{bmatrix}
\]

South

\[
\begin{bmatrix}
5 & -3 & -3 \\
5 & 0 & -3 \\
5 & -3 & -3 \\
\end{bmatrix}
\]

South east

\[
\begin{bmatrix}
-3 & -3 & -3 \\
-3 & 0 & -3 \\
5 & 5 & -3 \\
\end{bmatrix}
\]

East

\[
\begin{bmatrix}
-3 & -3 & -3 \\
-3 & 0 & -3 \\
5 & 5 & 5 \\
\end{bmatrix}
\]

North east

\[
\begin{bmatrix}
-3 & -3 & -3 \\
-3 & 0 & 5 \\
-3 & 5 & 5 \\
\end{bmatrix}
\]
IV. SECOND DERIVATIVE OPERATORS

These operators work on the sole principle of zero crossing, where they detect edges based on the fact that at the place in an image where the sign changes from negative to positive, the edge lies on the pixel containing the value zero.

Figure1: Depicts the use of second derivative operators for edge detection

A. Laplacian Edge Detector

In this technique we use a predetermined mask as in the case of other presented first derivative operators. This mask is convoluted over the complete image and the pixels containing the second derivative output as zero are marked. This technique is extremely useful when the image being analysed is void of any noise as the Laplacian edge detector is extremely sensitive to noise. [7][8]

B. Laplacian Of Gaussian Detector

As stated earlier the use of Laplacian operator is sensitive to images containing noise and to remove this inadequacy we use the LoG (Laplacian of Gaussian) method. In this technique at first a smooth filter similar to that of a Gaussian filter is applied so as to remove the effectively present noise in the image followed by convolution from the Laplacian Operator. Another way to implement this function is to pre-calculate the kernel by convoluting the Laplacian of Gaussian filters that need to applied on an image. [9]

C. Difference Of Gaussian Detector

This is a feature enhancement algorithm which removes the overall high frequency information from the image and tries to detect the edges by conforming to an output which is resulted from the difference of two images. These images are obtained by applying Gaussian filters of different intensities on them and then subtracting the two images to obtain the image with relevant information. In this case the edges. This method works similarly as that of band pass filter as the relevant information present between the frequencies associated with two frequencies of the Gaussian filters applied is only available.
D. Marr-Hildreth Algorithm

This algorithm works dependently on the working of LoG or DoG. In the LoG it operates by simply convolving the image with the Laplacian of the Gaussian function and in the case of DoG it works by calculating and finding out fast approximations presented by the output from the DoG. This algorithm has two major drawbacks. First is its ability to detect and present False edges; false edges are areas in an image containing no distinct edges but still being qualified as edges by the algorithm. Second is the localization error which is extremely high in case of curved edges [11].

V. MULTI-LEVEL THRESHOLDING AND OPTIMIZATION TECHNIQUES

Multi-Level Thresholding can be defined as separating the grey scale image into different thresholds depending on the characteristics of the image. The determination of the threshold vector is performed by applying optimization algorithms to derive the optimal solution. The main objective of the optimization algorithms is to improve the performance of the function used to find the threshold. We choose multi-level thresholding technique for image segmentation as it works best with complex images which contains various aspects to be analyzed where bi-level thresholding fails to work efficiently. There are many multi-level thresholding techniques available for image segmentation but we are going to focus on the entropy criterion method because it can be optimized using various optimization techniques efficiently. (Kapur Method) [10]

Kapur Method for Multilevel Thresholding:

This method was initially developed by Kapur; a method useful for bi-level thresholding which was described as follows:

We take the image to contain a maximum of L grayscale values or grey levels where \( f(m,n) \) represents the grey level of the image at the position \((m,n)\) in a MxN image. The value of \( f(m,n) \) is always present between the range \([0,L-1]\).

Now, let \( h_i \) represent the sum of all pixels containing grey scale values corresponding to \( i \). We can determine the joint Probability mass function as follows:

\[
p_i = \frac{h_i}{N} \tag{1}
\]

then our main objective will be to maximise the fitness function given by

\[
F(t) = H_0 + H_1 \tag{2}
\]

where \( H_0 \) and \( H_1 \) are given by:

\[
H_0 = \sum_{i=0}^{L-1} \frac{p_i}{w_0} \ln \left( \frac{p_i}{w_0} \right) \tag{3}
\]

\[
H_1 = \sum_{i=0}^{L-1} \frac{p_i}{w_1} \ln \left( \frac{p_i}{w_1} \right) \tag{4}
\]

\[
w_0 = \sum_{i=0}^{L-1} p_i \tag{5}
\]

\[
w_1 = \sum_{i=0}^{L-1} p_i \tag{6}
\]
This method for bi-level can be more generalised to contain N level of threshold as follows:

\[ F(t_1, t_2, \ldots, t_m) = H_0 + H_1 + H_2 + \ldots + H_m \]

where \( H_m \) is given by:

\[ H_m = \sum_{i=1}^{m} \frac{P_i}{W_m} \ln \left( \frac{P_i}{W_m} \right) \]

Next we apply optimization algorithms onto the Multi-level Thresholding Kapur algorithm to find and optimize the Threshold Vector. The algorithms being discussed here are: Particle Swarm Optimization algorithm, FireFly based Optimization Algorithm, Elephant Herding Algorithm

### A. Particle Swarm Optimization

The particle swarm optimization algorithm is derived from the observation of social behavior of the bird flock. It was developed by Kennedy and Eberhart in 1995. In PSO the potential solutions which in our case would be the values of the Threshold vector move around the problem space in search of a solution. [10]

Each particle in the problem space moves around to find optimum solution based on 3 essential features:

1. Based on its own experience.
2. Based on experience of nearby particles

The main function carried out by the optimisation algorithm is to carry out the computation of the velocity of the particle and the next position of the particle in the problem space. Let i denote a particular particle present in the swarm, then:

\[ v_{id}^{t+1} = w \cdot v_{id}^t + c_1 \cdot \psi_1 \cdot (p_{id}^t - x_{id}^t) + c_2 \cdot \psi_2 \cdot (p_{gd}^t - x_{id}^t) \]

\[ x_{id}^{t+1} = x_{id}^t + v_{id}^{t+1} \]

where \( x_{id}^{t+1} \) denotes the next position depending on current position \( x_{id}^t \) and \( v_{id}^{t+1} \)

\( v_{id}^{t+1} \) denotes new velocity depending upon three factors:

1. \( w \cdot V_i \) where \( W \) is the weight factor
2. \( c_1 \cdot r_1 \cdot (p_{bi} - x_i) \) where \( c_1 \) is self confidence factor , \( r_1 \) is random variable of range \([0,1]\) and \( p_{bi} \) is the personal best of i.
3. \( c_2 \cdot r_2 \cdot (g - x_i) \) where \( c_2 \) is the confidence vector of swarm , \( r_2 \) is another random variable in range \([0,1]\) and \( g \) is the global best of the swarm.

Now the algorithm for optimum solution of Kapur Multi-level Thresholding based on PSO is as follows:

**Step1:** initialize randomly the max and min limits of threshold values and set the iteration of t to zero. t represents the given threshold value.
Step 2: Evaluate values of the particles using the fitness function for Kapur method and then compare the personal best of each particle to its current fitness and set \( p_{b_i} \) to the better performance.

Step 3: Change the velocity vector for each according the PSO formula and move each particle to its new position.

Step 4: Go to Step 2 and repeat until it gets an optimum solution. [10]

B. Firefly Optimisation Algorithm

In the paper [11] Firefly algorithm was developed by Xin-she Yang at Cambridge University in 2008. In this algorithm we follow three basic rules: (1) All fireflies are unisex so they are attracted to each other regardless of their sex; (2) attractiveness is directly proportional to brightness, thus if their exist two fireflies having different brightness values the one with lower brightness will get attracted to the one with higher brightness; (3) If all the fireflies have same brightness than the firefly will move randomly. Here we select a monotonically decreasing function of the distance \( r_{ij} \), where \( r_{ij} = d(x_j, x_i) \) for the \( j^{th} \) firefly.

Here we apply an exponential function as follows:

\[
R(i, j) = \frac{\sum_{k=1}^{n} (X(i, k) - X(j, k))^2}{\beta_0 e^{-\gamma R(i, j)}} \tag{11}
\]

\[
\beta \leftarrow \beta_0 e^{-\gamma R(i, j)} \tag{12}
\]

Where the \( \beta_0 \) is the attractiveness at \( r_{ij} = 0 \) and \( \gamma \) is the light absorption coefficient at the source. The movement of a firefly \( i \) attracted to another firefly \( j \) where \( j \) is more attractive is given by:

\[
X(i, k) \leftarrow (1 - \beta)X(i, k) + \beta X(j, k) + u(i, k) \tag{13}
\]

\[
u(i, k) = \alpha (rand(1 - \frac{1}{2})) \tag{14}
\]

If there is no brighter firefly and all the fireflies have same level of brightness than the movement of any firefly will be random and firefly \( x_i \) will move randomly according to the following equation:

\[
X(i_{\_max}, k) \leftarrow X(i_{\_max}, k) + u(i_{\_max}, k), k = 1, 2, \ldots c \tag{15}
\]

\[
u(i_{\_max}, k) = \alpha (rand(2 - \frac{1}{2})) \tag{16}
\]

where \( rand1 = U(0, 1) \) and \( rand2 = U(0, 1) \) are random numbers obtained from the uniform distribution; The brightness of a firefly is determined by using MEFT (maximum entropy firefly thresholding) algorithm. Firefly algorithm is used to optimise the fitness function of Kapur.

C. Elephant Herding Algorithm

[12] Elephant herding optimization algorithm was proposed by Wang et. al. in 2016. This algorithm has simplified elephant herding behavior into three simple rules: Elephant population is composed of clans where each clan is headed by some matriarch or male member of the clan. There are some fixed numbers of elephants in each generation who leave their family and live far away from their clan. Elephant population is divided into \( k \) clans. Each clan is led y a matriarch \( c_i \). Each and every member \( i \) of clan \( j \) move according to the leader according to the equation:
where \( x_{\text{new, }i,j} \), \( c_i \), \( j \) represents new position of elephant \( j \) in clan \( i \). Here old position is given by \( x_{i,j} \) and the best solution or the matriarch of the clan is given by \( x_{\text{best, }i} \). \( \alpha \in [0,1] \) is a scale factor which determines the influence of the matriarch and \( r \in [0,1] \) is a random number which helps in improving diversity of the population. Position of the matriarch also the best elephant in the clan is updated using the following function.

\[
x(\text{new}, ci) = \alpha \cdot x(\text{center}, ci) + \frac{r}{\text{rand}} \cdot (x(\text{best, }ci) - x(ci, j))
\]  

where \( \beta \in [0,1] \) controls the influence of \( x_{\text{center},i} \) which is given as

\[
x(\text{center}, ci, d) = \frac{1}{N_c} \sum_{i=1}^{N_c} x(ci, l, d)
\]  

Male elephants that separate from the group are used to show exploration. In each generation and in each clan \( m_c \) elephants with the worst fitness values are moved according to the following equation:

\[
x(\text{worst, }ci) = x_{\text{min}} + x_{\text{max}} - x_{\text{min}} + 1 \cdot \text{rand}
\]  

where \( x_{\text{min}} \) and \( x_{\text{max}} \) represent lower and upper bound of search space and rand is a random number generated from uniform distribution.

**VI. PERFORMANCE EVALUATION CRITERIA FOR IMAGES**

The performance evaluation of edge detection algorithms is done here by detection of true edges, error ratio, and noise level here, the paper compares first derivative operators, second derivative operators and other evolutionary algorithms with mean square error, peak signal to noise ratio and root mean square error.

**A. MEAN SQUARE ERROR**

The MSE is the cumulative squared error between the compressed and the original image. It specifies the average difference between the original image and the edge detected image [4]. The higher value of MSE means greater difference between the detected image and the original image. Generally value of MSE is low when performing image analysis but, when performing edge detection value of MSE should be high because it ensures detection of weaker edges too.

\[
MSE = \frac{1}{MN} \sum_{y=1}^{M} \sum_{x=1}^{N} [I(x, y) - \Gamma(x, y)]^2
\]  

where \( I(x,y) \) is the original image, \( \Gamma(x,y) \) is the approximated version and \( M,N \) are the dimensions of the images

**B. PEAK SIGNAL TO NOISE RATIO**

Peak signal to noise ratio often known as PSNR is used to measure the peak error. It calculates ratio between maximum power of signal that is possible and power of the noise that corrupts the given image [4]. A higher value of PSNR shows that the image formed is of higher quality. But, like edge detection in some cases PSNR should be lesser to get appropriate results. It is calculated as:

\[
\text{PSNR} = 10 \log_{10} \frac{R^2}{\text{MSE}}
\]

\( R \) is the maximum possible variation in the input image. \( R \) is 255 if it has a data type which is 8 bit unsigned integer.
VII. EXPERIMENTAL RESULTS

The experiment is done to compare first derivative operators and second derivative operators. The test is based on performance of operators with respect to MSE, PSNR and RMSE. The picture taken for the experiment has resolution of 512 x 512.

Table 1 shows grey scale images and their corresponding edge detected outputs. Table 2 shows MSE, RMSE and PSNR values of different first derivative and second derivative operators and graph 1, graph 2 and graph 3 represents the MSE, RMSE and PSNR values in terms of bar graph respectively.

Figure 2: Original Image

Figure 3: Sobel detected image
Figure 4: Roberts detected image

Figure 5 Robinson detected image

Figure 6: Prewitt detected image
Figure 7: Kirsch detected image

Figure 8: Laplacian detected image

Figure 9: Laplacian of gaussian detected image
Figure 10: Difference of Gaussian detected image

<table>
<thead>
<tr>
<th>Operator</th>
<th>MSE</th>
<th>RMSE</th>
<th>PSNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sobel</td>
<td>1.77E+04</td>
<td>132.9322</td>
<td>5.6582</td>
</tr>
<tr>
<td>Prewitt</td>
<td>1.77e+04</td>
<td>132.9325</td>
<td>5.6582</td>
</tr>
<tr>
<td>Roberts</td>
<td>1.77e+04</td>
<td>132.9328</td>
<td>5.6582</td>
</tr>
<tr>
<td>Robinson</td>
<td>1.96e+04</td>
<td>139.9409</td>
<td>5.2119</td>
</tr>
<tr>
<td>Kirsch</td>
<td>1.98e+04</td>
<td>140.575</td>
<td>5.1726</td>
</tr>
<tr>
<td>Laplacian</td>
<td>1.87e+04</td>
<td>136.7023</td>
<td>5.4153</td>
</tr>
<tr>
<td>LoG</td>
<td>1.77e+04</td>
<td>132.9328</td>
<td>5.6711</td>
</tr>
<tr>
<td>DoG</td>
<td>1.76e+04</td>
<td>132.7351</td>
<td>5.6582</td>
</tr>
</tbody>
</table>

Table 1: Represents the MSE, RMSE and PSNR values for first and second derivative operators.

Graph 1: MSE values plotted corresponding to the method.
Graph 2: RMSE values plotted corresponding to the method.
Graph 3: PSNR values plotted corresponding to the method.
VIII. CONCLUSION

Various edge detectors have been studied and compared on the basis of various factors such as RMSE, PSNR, and MSE. The main aim of this paper is to present most efficient image segmentation algorithm which can detect edges with minimum errors and which can handle several noises more effectively. On the basis of MSE Kirsch Compass edge detection is efficient because higher value of MSE ensures detection of weaker edges. According to PSNR Kirsch is effective because lower value of PSNR provides appropriate results. In conclusion, the ideal image segmentation method is the one which can detect edges without eliminating important details. Various factors are involved for selection of a particular algorithm for an input image and based on such factors only one can decide the suitable algorithm to be used.

REFERENCES


[2]. Tzu-Heng Henry Lee“Edge Detection Analysis” in Graduate Institute of Communication Engineering, National Taiwan University, Taipei, Taiwan, ROC

[3]. Mr. Salem Saleh Al-amri1, Dr. N.V. Kalyankar2 and Dr.Khamitkar S.D” Image Segmentation by using edge detection” in (IJCSE) International Journal on Computer Science and Engineering, Vol. 02, No. 03, 2010


[9]. AshishAnand, Dr.Sanjaya Shankar Tripathy, Dr. R. Sukesh Kumar "An Improved Edge Detection Using Morphological Laplacian of Gaussian Operator” in 2nd International Conference on Signal Processing and Integrated Networks (SPIN) 2015
