Automatic Multi Stage Image Segmentation Using Normalized Cut in Gradient Image

Basavaprasad B.*
Research and Development Center,
Bharathiar University, Coimbatore-641 046, India.
(Corresponding Author*)

Ravindra S. Hegadi
Department of Computer Science
Solapur University, Solapur - 413255, India.

Abstract

Image segmentation is one of the main phases in the image processing which is processed for separation of image pixels into significant image regions. These regions are matched to objects, individual surfaces, and occlusion boundary estimation within motion or stereo systems, natural parts of objects which are used for object recognition, image compression or image editing. It is very challenging process to segment the different objects of the image. For that we have proposed a method for Automatic Multi Stage Image Segmentation Based on Normalized Cut in Gradient Image. There are three main steps in order to segment the image. During the first step, edge-preserving statistical noise reduction approach is used as a pre-processing stage in order to enhance the image. Then an accurate estimate of the image gradient is computed. In second step we separate the image into several regions, and then segments are produced locally from each of the image cells. Finally in the third stage segmentation and possible merging is carried out. As a result we get meaningful segments which are our required output.

Keywords: Gradient image; multi-stage; image segmentation; threshold value; normalized cut.
1. INTRODUCTION

Image segmentation has been used in various fields of image processing. Many of the computer vision applications need image segmentation for information extraction. The objective of image segmentation is to divide the source image into significant regions or segments that are associated with all kinds of images and these significant regions or segments are used for auxiliary analysis of images [3]. Image segmentation is very important phase of image processing which is used in many practical applications like object recognition, occlusion boundary estimation within motion or stereo systems, image compression or image editing. Extensive research has been done to develop numerous techniques to produce image segmentation which are very close to human’s observations. During the process of image segmentation the properties of image such as brightness, contrast, and edge clarity may be impaired due to noise and other matters. Hence it is important to enhance the image and preserve the edges of the image before it is being segmented [1]. We propose new multi stage image segmentation based on normalized cut in gradient image to resolve the above stated problem. Initially the input image is pre-processed. This process consists of two phases. During the first phase noise of the input image is removed using Guided filtering. Then the image is processed under guided filtering for edge preserving. Finally multi stage normalized image segmentation is applied to get better segmentation results.

We organized our paper as follows. The section 2 explains the survey that reveals the related work. Proposed method is presented in section 3. Experimental results are discussed in section 4 and finally conclusions are given in section 5.

2. LITERATURE SURVEY

In this section we discuss the related work on the image enhancing and segmentation techniques.

2.1. Edge detection and Preserving (Image Enhancement)

Noise removal or cancellation can result in poor quality of image such as low contrast, unclear edges and blur. To overcome these problems, image enhancement techniques are used. The process of image enhancement can recover the clarity of image for human perception. This technique essentially provides improved input for further automated image processing particularly image segmentation methods. The core objective of image enhancement is to alter the properties of an image which is further suitable for image segmentation. The selection of one or more properties of an image during the modification is much perceived to a given task. Likewise depending on observer’s practice and the visual system, particular image enhancement techniques are selected.

Removing blurring, increasing contrast and illuminating are examples of enhancement operations. Eliminating blurring and adding the range of contrast will enhance the quality of image. The original image may contain very high and very low intensity
values, which uncovers the details of an image. An adaptive algorithm with enhancement can disclose these details. Algorithms which are adaptive change their procedure depending on the pixel and other information of an image region under processing. Many techniques have been developed which can enhance a digital image without the loss of image content. The techniques of image enhancement can roughly be divided into the following two types [4].

1. Spatial Domain Methods
2. Frequency Domain Methods

In the first type that is in spatial domain methods the image pixels are directly taken as criteria for image enhancement. The values of the pixel are manipulated in order to achieve the required image enhancement. In the second type that is frequency domain techniques, initially the image is transmitted into frequency domain. That is Fourier Transform of the image is computed first. Later every other enhancement operations are carried out on the computed Fourier transform of the image. Finally the Inverse Fourier transform is calculated to get the required image. The image properties such as contrast, brightness and the grey level distribution are modified using these enhancement operations. As a result the intensity or the pixel values of the output image will be modified according to the application of transformation function on the input pixel values. The applications of image enhancement exist in every field where understanding of images is must and also in the analysis of images for example, computer vision, medical image analysis, traffic image analysis, remote object detection, satellite image analysis etc.

2.2. Image segmentation using normalized cuts

J. Shi and J. Malik proposed Normalized cuts for image segmentation. In their paper, they proposed a novel method for solving the problem of image segmentation. Instead of aiming on local features and their reliabilities in the image information, they advanced at mining the global features of an image. Here the image segmentation problem is treated as a problem of graph partitioning. They proposed global features criteria which are normalized cut for segmentation of the graph. This normalized cut technique computes together the total dissimilarity between different segments or clusters and also the total similarity within each segment. They showed that proficient computational method grounded on a comprehensive eigen-value problem can be used to optimize the benchmark [8].

M.Y. Choong, C.F. Liau, J. Mountstephens, M.S. Arifianto, and K.T.K. Teo proposed multistage image clustering and image Segmentation using Normalized Cuts [5]. This technique needs huge computation of similarity dimension for the segmentation of an image. As the images taken from a digital camera are of high resolution, they can be resized to a resolution so that the algorithm can perform segmentation with less effort. They divided an image into the same size of regions or segments, also called as image cells, for the segmentation. They solved the problem of
significant properties, which are missed when the image resolution is overly reduced. Later on, these locally segmented segments or clusters from the image cells are given for the second stage process of segmentation and then global merging is applied to them.

Subhransu Maji, Nisheeth K. Vishnoi, and Jitendra Malik proposed a method for biased normalized cuts [7]. They modified the conventional normalized cuts to integrate priors which are used for constrained image segmentation. In comparison with previous normalized cuts methods which integrate constraints, this method has two advantages. First, looking for the solutions that are adequately correlated with priors, which are permitted using noisy top-down information. Additional, advantage is the enhancement in the excellence of image segmentation.

3. PROPOSED WORK

In the proposed work a new evolutionary method is introduced which contains effective steps to get better segmentation Results.

The algorithm is as follows.

Algorithm 1:

1. Read input color image to be segmented
2. Perform edge preserving using Guided Filtering
3. Construct an $M \times N = n$ image cells for partition an input gradient image
4. calculate 1st image cells generation that is required to segment
   
   If $\delta > \text{threshold value}$
   
   Proceed to local normalized cuts segmentation

   else
   
   Calculate background node in first image cell

   end
5. Initial stage segmentation
   
   a. Normalized cuts local segmentation into $k_1$ segments on first image cell
      until last image cell.

   b. Every segmented image cells is represented by a node.
6. Next stage segmentation
   
   Implement Normalized cuts segmentation based on the figured nodes to $k_2$ segments.
7. Output: Produce segmented image.
3.1 Preprocessing

Segmentation of image and consequential cropping from its background, which is affected by noise, is still remained a complex job in the area of image processing. Therefore the input image is preprocessed before it is segmented.

Guided Filtering:

The guided filter is defined as a local linear representation along with the guidance \( I \) and the filtering output \( q \). It is implicit that \( q \) is a linear transmute of \( I \) in a window \( W_k \) centered at the pixel \( k \).

\[
q_i = a_k I_i + b_k, \forall i \in \omega_k
\]

(1)

Where \((a_k, b_k)\) are some linear coefficients understood to be constant in \( W_k \). A square window of a radius \( r \) is used. This linear model, which is local, make sure that \( q \) has an edge if and only if \( I \) has an edge, for the reason \( \nabla q = a \nabla I \).

To control the linear coefficients \((a_k, b_k)\) we need constraints from the filtering input \( p \). We model the output \( q \) as the input \( p \) by subtracting some unwanted constituents \( n \) like textures or noises.

\[
q_i = p_i - n_i
\]

(2)

A solution is obtained that minimizes the difference between \( p \) and \( q \) at the same time preserving the linear model. Particularly, the following cost function is reduced in the window \( \omega_k \):

\[
E(a_k, b_k) = \sum_{i \in \omega_k} \left( (a_k I_i + b_k - p_i)^2 + \epsilon k^2 \right)
\]

(3)

Here, \( \epsilon \) is a regularization parameter penalizing large \( a_k \). The model for linear ridge regression and its solution is given by,

\[
a_k = \frac{1}{|\omega_k|} \sum_{i \in \omega_k} I_i p_i - \mu_k \bar{p}_k}{\sigma_k^2 + \epsilon}
\]

(4)

\[
b_k = \bar{p}_k - a_k \mu_k
\]

(5)

Where, \( \mu_k \) and \( \sigma_k^2 \) are the means and variance of \( I \) is \( w_k \), \( |w_k| \) is the pixel number in \( w_k \), and \( \bar{p}_k = \frac{1}{|w_k|} \sum_{i \in w_k} p_i \) is the mean of \( p \) in \( w_k \). Having obtained the linear coefficients \((a_k, b_k)\), we can compute the filtering output \( q_i \) by using Equation (1).
Nevertheless, a pixel $i$ is sophisticated in all the overlying windows $w_k$ that covers $i$, so the value of $q_i$ in Equation (1) is not matching when it is figured in different windows. Calculating the average of all the possible values of $q_i$ is the simple tactic. Therefore following the computation of $(a_k, b_k)$ for $w_k$ every window in the image, the computation of the filtering is given by the formula

$$a_k = \frac{1}{|\omega_k|} \sum_{k \in \omega_k} a_k I, b_k$$

Noticing that $\sum_{k \in \omega_k} a_k = \sum_{k \in \omega_k} a_k$ due to the symmetry of the box window, we rewrite Equation (2) as

$$q_i - \bar{a}_i I_i + \bar{b}_i$$

Where $\bar{a}_i = \frac{1}{|\omega|} \sum_{k \in \omega} a_k$ and $\bar{b}_i = \frac{1}{|\omega|} \sum_{k \in \omega} b_k$ are the average factors of every windows overlapping $i$. $\nabla q$ is no longer scaling of $\nabla I$ because the linear coefficients $(\bar{a}_i, \bar{b}_i)$ vary spatially. But as $(\bar{a}_i, \bar{b}_i)$ are the outputs of a mean filter, their gradients can be predictable to be much lesser than that of $I$ near clear edges. In this circumstances still we can have $\nabla q \approx \bar{a} \nabla I$, it means that unexpected intensity modification in $I$ can be typically preserved in Equations (4), (6) and (7), which corresponds to the guided filter definition. A pseudo-code for guided filtering is shown in the following algorithm. In this algorithm, $f_{\text{mean}}$ represents mean filter with a window radius $r$. The abbreviations of correspondence ($\text{Corr}$), variance ($\text{Var}$), and covariance ($\text{Cov}$) point to the sensitive meaning of these variables.

**Extension to Color Filtering**

The guided filter can be easily prolonged to color images. In the case when the filtering input $p$ is multichannel, it is forthright to apply the filter to each channel self-reliantly. In the case when the guidance image $I$ is multichannel, we rewrite the local linear model as

$$q_i = a_k^T I_i + b_k, \forall i \in \omega_k$$

Here $I_i$ is a $3 \times 1$ color vector, $a_k$ is a $3 \times 1$ coefficient vector, $q_i$ and $b_k$ are scalars. The guided filter for color guidance images becomes

$$a_k = (\Sigma_k + \varepsilon U)^{-1} \left( \frac{1}{|\omega_k|} \sum_{i \in \omega_k} I_i p_i - \mu_k \bar{P}_k \right)$$
Here, $\Sigma_k$ is the $3 \times 3$ covariance matrix of $I$ in $W_k$, and $U$ is a $3 \times 3$ identity matrix. A color guidance image can better preserve the edges that are not discrete in gray-scale. This is also the case in bilateral filtering [6].

After the preprocessing step the image is free from noise as well as the image is enhanced so that that this image is very feasible to segment.

4. IMAGE SEGMENTATION

It is not likely that an image to be segmented in accordance to human perspective, when the segmentation method is based on original information, such as image brightness [9]. Thus, the aim of the segmentation should be focused on locating objects and boundary so that the image can be meaningful and easier to examine.

4.1 Image Segmentation using Normalized Cuts

Here the filtered image segmented using three different attributes first the texture features, second color values (red, blue and green) and spatial data of an image.

4.1.1 Extraction of texture features

Before the image is segmented we extract four texture features arithmetic mean, variance, skewness and kurtosis which are explained as follows.

*Arithmetic Mean*

The arithmetic mean also known as averaging filter operated on ‘$m \times n$’ window is calculated by averaging all pixel values inside the window and substituting the center pixel value in the endpoint image with the result. Its mathematical formulation is as follows.

$$f(x, y) = \frac{1}{mn} \sum_{(r,c) \in W} g(r,c)$$

(12)

Where, $g$ is the noisy image and $r$ and $c$ are row and column respectively, within a window $W$ of size $m \times n$, where the operation takes place. In Figure 2 we can observe the arithmetic mean of the input image.
Variance

The variance is defined as a measure of how remote the group of numbers is range out. It is one of numerous descriptors of a distribution probability, telling how remote the numbers exists from the expected value i.e., mean. In specific, the variance is defined as the instants of a distribution. In that regard, it forms portion of a methodical approach to differentiate between probability distributions. Although other such methodologies have been developed, those based on instants are beneficial in terms of scientific and computational simplicity. Mathematically variance is defined as

\[
\tilde{f}(x, y) = \frac{1}{mn-1} \sum_{(r,c) \in W} \left( g(r,c) - \frac{1}{mn-1} \sum_{(r,c)} g(r,c) \right)^2
\]

(13)

Variance can be used to determine the edge position in image processing. In Figure 2 we can observe the variance of the input image.

Skewness

In statistics, skewness is defined as a degree of the irregularity of the distribution probability of a value in real random variable. The skewness value can be either positive or negative, or may even be indeterminate. Qualitatively, a negative skew indicates that the tail on the left side of the probability thickness function is lengthier than the right side and the bulk of the values perhaps comprising the median lies at the right of the mean. The zero value stipulates that the values are moderately, consistently distributed on both sides of the mean, characteristically but not essentially suggesting a symmetric distribution. Mathematically skewness can be given by

\[
\hat{f}(x, y) = \frac{1}{mn-1} \sum_{(r,c) \in W} \left[ \frac{1}{mn-1} \sum_{(r,c) \in W} \left( g(r,c) - \frac{1}{mn-1} \sum_{(r,c)} g(r,c) \right)^3 \right]^{1/2}
\]

(14)

We can use skewness in making judgments about image surfaces. In Figure 2 we can observe the skewness of the input image.

Kurtosis

In statistics, kurtosis is a degree of the contour of the probability circulation of a real-valued random variable. It is closely related to the quarter moment of a circulation. A high kurtosis circulation has longer, fatter tails, and often (but not always) a sharper peak. A low kurtosis circulation has shorter, thinner tails, and often (but not always) a more rounded peak. Mathematically kurtosis is given as follows
In digital image processing, kurtosis values are interpreted in combination with noise and resolution measurement. High kurtosis values should go hand in hand with low noise and low resolution. In Figure 2, we can observe the kurtosis of the input image.

In this paper, the image segmentation method used here is based on graph theoretic approach which generally performs pixels grouping into regions and it is treated as graph partitioning problem \[9\]. A set of image pixels can be represented as weighted graph \( G = (V, E) \), where \( V \) represents the vertices, which are image pixels (a vertex is made up of one node) and \( E \) represents the edges in the form of weights, \( w \). Each of the \( w \) gives a measurement of the correspondence between node \( i \) and node \( j \). A graph is bi-partitioned into two distinct sub-graphs \( A \) and \( B \) with the form that it reduces the value of

\[
cut(A, B) = \sum_{i \in A, j \in B} w(i, j)
\]

(16)

Where \( A \bigcup B = V \) and \( A \bigcap B = \emptyset \) being the constraints.

The degree of the dissimilarity between two sub-graphs \( A \) and \( B \), which is the sum of weights of the pairs of nodes, are to be removed. The bi-separating process recursively finds the least cuts until a number of \( k \) sub-graphs are formed, with the condition that the maximum possible cut across the sub-graphs is minimized. In other words, the formed distinct sub-graphs have high similarity within the sub-graphs and low similarity across different sub-graphs. However, the least cut principle falls short that its cut algorithm favors in cutting isolated nodes to form sub-graphs. This lead to proposing another cut algorithm based on regularized cut (Ncut) principle to alleviate the problem. The regularized cut criterion is derived as in

\[
Ncut(A, B) = \frac{\text{cut}(A, B)}{\text{assoc}(A, V)} - \frac{\text{cut}(A, B)}{\text{assoc}(B, V)}
\]

(17)

Where,

\[
\text{assoc}(A, V) = \sum_{i \in A, j \in V} w(i, j)
\]

The \( Ncut(A, B) \) is then converted to

\[
(D - W) = \lambda D y
\]

(18)
for solving the eigenvectors $y$ and eigenvalues $\lambda$. $D - W$ is called the Laplacian matrix whereby $W$ is a symmetrical matrix with $W(i,j) = w(i,j)$. Each of the weight elements in matrix $W$ is defined as

$$W(i,j) = e^{-\frac{(F(i) - F(j))^2}{\sigma_f}} \times \begin{cases} e^{-\frac{(X(i) - X(j))^2}{\sigma_X^2}} < r \\ 0, \text{otherwise} \end{cases}$$  \hfill (19)

Where $F(i)$ feature vector is based on intensity value in color of node $i$ and $X(i)$ is the spatial location of the node. When a pair of node $i$ and $j$ is more than $r$ number of pixels apart, the weight $w(i,j)$ is considered 0. $D$ is a diagonal matrix with $d(i) = \sum_{j=1}^{M} w(i,j)$ on its diagonal. An Image segmentation is then done by clustering based on the eigenvector. Eigenvector corresponds to the second smallest eigenvalue chosen for a clustering algorithm to partition the image. As soon as clustering process completed, the required segmentation result is finally achieved which is grounded on the clusters which are formed amongst the eigenvector.

### 4.2 Image Cells Generation for Image Subdivision

Normalized cuts implementation in image subdivision has a disadvantage in its calculation. The image of size $m \times n$ should undergo normalized cuts using the $W$ matrix with a size of $(m \times n) \times (m \times n)$. The image contains different resolution in itself. Therefore, an adjacency matrix is measured as a very big matrix in order to solve to solve for eigenvalues using the computer.

**Image Cells Generation and Segmentation in Image Cells**

Processing of this large matrix should be avoided. Therefore the high resolution image should be divided into equal sized image cells. Subsequently natural and generic image has uneven pixels distribution such that the occurrence of image pixels for specific range occurred unequally throughout the entire area of an image. The first stage of segmentation is begun by providing the $k_i$ number of clusters, normalized cuts algorithm is later achieved for division out $k_i$ number of clusters for the particular cell. During this segmentation phase, over-segmentation could occur as perception power in an image cell is reduced compared with segmentation on a entire image directly. However, it aids to reduce the tendency of object boundaries which are missing. The segmentation process is repeated for each cells and is done independently. These segmented clusters are then given for second stage segmentation.
Clusters Merging and Second Stage Segmentation

The second stage segmentation starts with the computation of simple representations of the clusters. Each segmented cluster has a node comprising the median value of the cluster and centroid of the cluster. The computed nodes are then used for second stage segmentation using normalized cuts algorithm. The value of threshold $r$, which is the maximum distance for a pair of pixels can be increased because the computed nodes are scattered in sparse manner. In addition, the standard deviation for the spatial location of the node, $\sigma$, has to be increased due to the matter also.

The number of nodes represents the overall segments from the first stage segmentation. For instance, 16 clusters from the first stage produce 16 nodes. The calculated nodes thereafter taken for parallel number using regularized cuts algorithm. These clusters then will be combined collectively and shared universal similarity depending on spatial and color location. A pseudo-code for normalized cuts with the execution of division of image cells is shown below.

Algorithm 2: Normalized cuts with image cells division

Input: Color/ grayscale image

1. Read input image
2. Construct $M \times N = n$ image cells
3. Complete statistical analysis of all image cells
4. Determine whether $i^{th}$ image cell required segmentation
   If $\delta >$ threshold value
   Proceed to local normalized cuts segmentation
   else
   Compute background node in $i^{th}$ image cell
5. First stage segmentation
   Normalized cuts local segmentation into $k_1$ clusters on $i^{th}$ image cell until $n^{th}$ image cell.
6. Computation of Nodes
   All of the clusters which are segmented are represented by nodes.
7. Second stage segmentation
   Perform normalized cuts on the nodes to segment $k_2$ clusters.

Output: Produce segmented clusters and display result

5. EXPERIMENTAL RESULTS AND DISCUSSION
The generic image segmentation using the proposed method is shown in Figure 2. We have obtained images of size $100 \times 70$ pixels for our experiment. We have tested the proposed algorithm using 200 images in our experiment. We have used Matlab 2014a software for our experiment. The average segmentation time is 5.2 seconds which is very fast. First the original input image is filtered using guided filtering. Then the image gradient is calculated. Then the texture features of image are extracted. The texture features are Mean, Variance, Skewness, and Kurtosis. Before the image is processed under normalized cut its color attributes red, blue and green are calculated. Finally the image is processed under normalized cut to get the final segmentation results.

**Figure 1**: Flowchart of the proposed algorithm
Figure 2: Image segmentation using the proposed method.
Table 1: Comparison between different graph cut functions

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Function</th>
<th>Optimization technique</th>
<th>Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimal cut [10]</td>
<td>$\text{Mincut} (A, B) = \sum_{u,v \in B} w(u,v)$</td>
<td>Gomory-Hu’s K-ways maxflow</td>
<td>Polynomial</td>
</tr>
<tr>
<td>Mean cut [11]</td>
<td>$\text{Meancut} (A, B) = \frac{\text{cut}((A, B) \setminus w(u,v))}{\text{cut}(A, B)}$</td>
<td>Minimum weight perfect matching</td>
<td>Polynomial time</td>
</tr>
<tr>
<td>Ratio cut [12]</td>
<td>$\text{Rcut} (A, B) = \frac{\text{cut1}(A, B)}{\text{cut2}(A, B)}$</td>
<td>Baseline</td>
<td>$O(n^{7/4})$</td>
</tr>
<tr>
<td>Proposed method</td>
<td>$\text{Ncut}(A, B) = \frac{\text{cut}(A, B)}{\text{vol}(A)} + \frac{\text{cut}(A, B)}{\text{vol}(B)}$</td>
<td>Generalized Eigen system</td>
<td>$O(mn)$</td>
</tr>
</tbody>
</table>

6. CONCLUSIONS

We have proposed a technique on Automatic Multi Stage Image Segmentation Based on Normalized Cut in Gradient Image. Our work is divided into three main steps. During the first step, the image is preprocessed by removing the noise present in the image using normalized least mean square method (NLMS) and then it is enhanced using the edge preserving method called guided filtering. In second step we separate the image into several regions, and then sectors are produced locally from each of the image cells. Finally, these segments then given for third and last step in which segmentation is done using normalized cuts to produce the meaningful segments. This will resulting in the number of image cells to be formed can be adaptively defined according to the image content. These assistances to rapidity of normalized cuts algorithm and matching the trade-off between efficiency of computation and proficiency such that the segmentation can see the ideal performance.

REFERENCES


