REGION-BASED IMAGE SEGMENTATION VIA GRAPH CUTS

Cevahir Çığla and A. Aydın Alatan

Department of Electrical and Electronics Engineering, M.E.T.U, Turkey
e-mail: cevahir@eee.metu.edu.tr, alatan@eee.metu.edu.tr

ABSTRACT

A graph theoretic color image segmentation algorithm is proposed, in which the popular normalized cuts image segmentation method is improved with modifications on its graph structure. The image is represented by a weighted undirected graph, whose nodes correspond to over-segmented regions, instead of pixels, that decreases the complexity of the overall algorithm. In addition, the link weights between the nodes are calculated through the intensity similarities of the neighboring regions. The irregular distribution of the nodes, as a result of such a modification, causes a bias towards combining regions with high number of links. This bias is removed by limiting the number of links for each node. Finally, segmentation is achieved by bipartitioning the graph recursively according to the minimization of the normalized cut measure. The simulation results indicate that the proposed segmentation scheme performs quite faster than the traditional normalized cut methods, as well as yielding better segmentation results due to its region-based representation.

Index Terms— Over segmentation, normalized cuts, color segmentation

1. INTRODUCTION

One of the most challenging problems in computer vision is color segmentation, which is also an irreplaceable tool for image understanding. Although many different approaches have been proposed for its solution, the problem has not been (and still far from being) completely solved due to its complicated, as well as subjective, range space. In the literature, most of the research has focused on solutions based on local properties [1][2][3]. Although they are quite efficient in general, the global characteristics might be lost during segmentation. On the other hand, within the last decade, graph theoretic segmentation methods have gained popularity, while utilizing graph cuts as their global optimization technique [4]-[10].

The graph theoretic approach introduces a top-to-bottom segmentation scheme in contrast to the local methods [4]. The image is considered as a “big picture” and mapped onto a weighted graph $G$, whose nodes ($V$) correspond to the pixels and the links ($E$) between the nodes are based on the similarities between these pixels. In different methods [4][5][7][9], segmentation is achieved by recursively bipartitioning the graph via minimizing cut measures according to the eigenvector decomposition of some special matrices. In [10], however, the partitioning is performed by utilizing nested cuts. A stochastic segmentation algorithm is also proposed in [8], based on $k$-way cuts.

It should also be noted that in global methods, the clustering is performed by a similar mechanism as in human perception, from the whole to the details. Although the global methods take advantage of this realistic assumption, their execution can be time-consuming as the image size increases. The computation problem is usually tried to be alleviated by only down sampling the images [4] or performing a multi-scale approach [7]. Unfortunately, the visual details of the images are usually lost during these operations.

In the proposed method, the normalized cuts image segmentation algorithm [4] is improved in order to decrease its computation time and increase its performance while preserving the visual details. To this aim, the graph structure is modified by assigning over-segmented regions as the nodes of the graph, instead of pixels, in order to decrease the graph complexity. Such a modification results in an irregular node distribution creating a bias towards combining regions with higher number of links. A solution to this problem is also proposed in order to remove the drawbacks of such a bias on the segmentation quality.

This paper is organized as follows; in Section 2 the overview of the normalized cuts image segmentation algorithm is given. The details of the proposed method are explained in Section 3. After presenting the experimental results in Section 4, the final section is devoted to the concluding remarks.

2. OVERVIEW OF NORMALIZED CUTS IMAGE SEGMENTATION

Normalized Cut Image Segmentation (NCIS) [4] is a global graph-based segmentation method that utilizes a splitting process beginning from the whole picture to the bottom. There are two main steps in NCIS, construction of the graph and iterative partitioning. The top-down property of the normalized cut method is provided by initially mapping the image onto a graph that holds the relations between pixels. An undirected weighted graph is constructed, in which the vertices correspond to the pixels and the link weights are evaluated via a linking cost function, given below.

$$w_{i,j} = \begin{cases} e^{-\frac{1}{2}(I(i)-I(j))^2/\sigma^2}, & |X(i) - X(j)|^2 < R \\ 0 & \text{elsewhere} \end{cases}$$

(1)

In (1), $I$ is the intensity image, $I(i)$ indicates the intensity value of the $i$th pixel and $X$ represents the location of these pixels. In the graph, the pixels which are located within a circle of radius $R$ are linked to each other; hence the graph is partial and the link weights define similarities between nodes as a function whose range space is $[0,1]$. 
The segmentation is achieved by division of the graph into smaller graphs by the cuts among the links, iteratively. At each step, one graph \((V)\) is partitioned into two sub-graphs, \(A\) and \(B\), such that \(V=\text{AUB}, \text{A}\cap\text{B} = \emptyset\). A cut on a link provides the separation of the two nodes connected to each other with the corresponding link. The cuts generally occur on more than one link and the separation of a set of nodes from another node group is provided, as in Figure 1. In addition, every cut has a cost value evaluated by the summation of the link weights belonging to the removed (cut) links, as follows

\[
\text{Cut} (A, B) = \sum_{i \neq A, j \neq B} W(i, j) \quad \text{(2)}
\]

NCIS method aims to minimize a normalized cuts measure (3) during the bipartitioning of a graph into \(A\) and \(B\).

\[
\text{Ncut} (A, B) = \frac{\text{Cut} (A, B)}{\text{TotalW} (A, V)} + \frac{\text{Cut} (A, B)}{\text{TotalW} (B, V)} \quad \text{(3)}
\]

\[
\text{Figure 1: A sample cut on a graph}
\]

In (3), \(\text{TotalW}(A, V)\) indicates the total link weight between the vertices in \(A\) and the whole graph \(V\). The minimization of the normalized cut exactly is an \(n\text{-complete}\) problem; however, an approximate discrete solution can be obtained by formulating the problem into real value domain and using change of variables [4].

The new formulation is given as:

\[
\text{Ncut} = \frac{\text{y}^T (D - W) \text{y}}{\text{y}^T D \text{y}} \quad \text{(4)}
\]

where \(D\) is an \(N \times N\) diagonal matrix of a graph with \(N\) nodes, indicating the total link weights belonging to each node individually, \(W\) is the affinity matrix of the graph showing the similarities between the nodes in a symmetric \(N \times N\) matrix. Finally, \(\text{y}\) is an \(N \times 1\) matrix, which is composed of real valued elements corresponding to the similarities of the nodes. The minimization for the formulation in (4) can be approximately achieved by the solution of the generalized eigenvalue system [4], as below

\[
D^{-1/2} (D - W) D^{-1/2} \text{y} = \hat{\lambda} \text{y} \quad \text{(5)}
\]

The eigenvalue decomposition of such a matrix is obviously time-consuming. Moreover, as the image size increases, the execution time of the algorithm will become longer and the memory requirements might cause some problems. One solution to these problems is down-sampling [4]. Unfortunately, the down-sampling operation causes the local intensity information to be lost and the object boundaries to be distorted in the final segmentation.

In this work, a modified normalized cuts algorithm is proposed in order to overcome this complexity problem without losing the local intensity information observed during down-sampling. The graph is constructed from the regions which are obtained by over-segmenting the original image via any local segmentation method. In this manner, the local properties are still embedded into these regions, while the number of the vertices in the graph is decreased to hundreds. The term over-segmentation is being used to describe any segmentation algorithm yielding relatively higher number of regions, which mostly do not have any semantic meaning, but correctly defining homogenous regions in terms of their color. The new graph structure (for the red squared region) is illustrated with an example in Figure 2. The random colored image corresponds to the over-segmented regions in the square and the centroids of these regions are utilized as the nodes of the graph.

Each node in the graph represents a small group of pixels, \(S_i\), and the mean intensity of \(S_i\) \((\bar{I})\) is utilized as the characteristic of the node, although other higher order descriptions, such as moments or histograms, could also be utilized. The link weights are calculated via the linking cost function below through \(S_i\)'s without considering the spatial locations of these segments.

\[
W_{i,j} = \begin{cases} 
\exp \left( \frac{\| \bar{I}_i - \bar{I}_j \|^2}{\sigma_i} \right) & \text{if } (i,j) \in E \\
0 & \text{elsewhere}
\end{cases} \quad \text{(6)}
\]

The new representation of the graph structure mostly causes an irregular distribution of the nodes among the graph. The reason of such a node distribution is due to the different sizes of \(S_i\)'s, as observed in Figure 2.

\[
\text{Figure 2: The modified graph structure}
\]

The irregular distribution of \(S_i\)'s on the graph introduces a tendency to group the regions having more connections together without checking the similarities in between. This tendency is
based on the formulation of the NCIS: when the minimization problem is analyzed, it could be observed that the normalization parameter $TotalW(A,V)$ is dependent on the total link cost, which is obtained for each node by the summation of equal number of link weights for the regular graph. Therefore, each link weight affects the partitioning process with the same ratio for the conventional pixel-wise application of NCIS. In the irregular distribution however, since the number of links for each node differ, the nodes with higher number of links have more terms in the summation and will result in a higher confidence, although the link weights are not strong enough. This phenomenon surpasses the weak links over the strong links, if they belong to a node having higher number of links. Moreover, such a tendency causes instabilities in the decomposition of the eigenvectors, which results in erroneous segmentation, although the NC is minimized. Considering the irregularities and different number of links for each node, during the minimization of the normalized cuts, the cut values should also be normalized with the number of the links ($N_A$, $N_B$). Hence the new minimization should be performed according to the following formula:

$$N_{cut}(A,B) = \frac{Cut(A,B)}{TotalW(A,V)/N_A} + \frac{Cut(B,A)}{TotalW(B,V)/N_B}$$

$$= \frac{N_A \cdot Cut(A,B)}{TotalW(A,V)} + \frac{N_B \cdot Cut(B,A)}{TotalW(B,V)}$$

(7)

However, the closed form solution to the new formulation can not be obtained by any change of variables. Hence, an approximate solution approach is presented by enforcing each node to have the same number of links. However, since the graph is undirected and the affinity matrix is symmetric, it is not straightforward, if not possible, to implement such a solution in practice. The equalization of links can be approximated by limiting the number of links for each node to a predetermined value and allowing less variation on the distribution of the number of links. Such a limitation provides nodes to have similar number of links, which can remove the suppression on strong links belonging to nodes with less links. Although the solution does not force nodes to have same number of links, the simulation results show that it dramatically decreases the irregularity.

4. EXPERIMENTS

In order to assess the performance of the proposed method; experiments have been conducted on different sets of images. In these experiments, the proposed method is compared with the traditional NCIS algorithm [4] and the multi-resolution normalized cuts [7] in terms of the segmentation quality and the execution time. The segmentation results are compared through mean-square error (MSE) with respect to the ground truth segmentation, as an objective comparison. MSE is evaluated as follows:

$$MSE = \frac{1}{m \times n} \sum_{i=1}^{m \times n} [G(i) - G^k(i)]^2$$

(8)

where $G$ is the ground truth segmentation mask of the image and $G^k$ is the segmentation mask of the $k^{th}$ method.

In Figures 3, 4 and 5 the segmentation results of the methods for different images (size 200x200) are illustrated. Based on the subjective comparison, the proposed method gives the most similar segmentation results to the ground truth for all the tested images. The resulting MSE plots for these images are also given in Figure 6. The proposed method gives the minimum MSE for the images Mosaic and Objects, on the other hand normalized cuts gives the minimum MSE for Cow. In addition, the execution time of the algorithms during the eigenvalue decomposition process after affinity matrix calculation is also given in Table 1. The experiments are performed by a 3 GHz CPU (4) Pentium computer and the codes are written in MATLAB. According to Table 1, the proposed method drastically decreases the execution time, as well as increasing the segmentation reliability in both subjective and objective criteria. During the experiments, the images are initially over-segmented into 500-700 sub-regions via Mean-Shift color segmentation [3] for the proposed method.

Finally, the effect of limiting the number of vertices for the nodes of the modified graph is presented in Figure 7 and Figure 8. In Figure 7, the values of the eigenvectors corresponding to second and third smallest eigenvalues are plotted pixel-wise for the two cases, limiting the number of links and un-limiting. In the eigenvector plots, it is expected to obtain similar values for the regions belonging to the same object. The object boundaries can easily be detected for the eigenvector plot of the limited link number case, while for the un-limiting case object boundaries are lost. The segmentation results of these eigenvector distributions are also given in Figure 8. As can be observed, the image is segmented erroneously due to the irregular distribution of the nodes; however the limitation corrects the segmentation.
Figure 5: (a) Cow image and (b) its ground truth segmentation, (c) NCIS, (d) proposed method, (e) Multi-resolution NC.

Table 1: The execution time comparison during the eigenvalue decomposition of the three methods

<table>
<thead>
<tr>
<th></th>
<th>Mosaics</th>
<th>Objects</th>
<th>Cow</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCIS</td>
<td>26,4</td>
<td>24,4</td>
<td>20,8</td>
</tr>
<tr>
<td>Proposed</td>
<td>2,05</td>
<td>1,41</td>
<td>2,24</td>
</tr>
<tr>
<td>Multi-res. NC</td>
<td>8,14</td>
<td>12</td>
<td>8</td>
</tr>
</tbody>
</table>

Figure 6: The MSE plots for the images with respect to the segmentation methods.

Figure 7: (a), (b) the second and the third smallest eigenvector of the Mosaic image with the number of links limited; (c), (d) the eigenvectors for the unlimited case.

Figure 8: The segmentation results of the Mosaics for the case with (a) unlimited link and (b) limited link

5. CONCLUSION

Increase in the performance of segmentation can be clearly observed throughout experiments. Utilization of over-segments, instead of pixels, decrease complexity of the graph structure and provides an initial step to group similarly colored pixels in local area. Thus, local regions are handled without any information loss while computation complexity is significantly decreased. The proposed method yields an improvement over the popular NCIS algorithm.

6. ACKNOWLEDGEMENT

This work is funded by TUBITAK under COST 292 Project.

7. REFERENCES