HiSeg: Unfolding of Segment Hierarchies in Color Images

Ahmad Ali Abin\textsuperscript{1} Farzane Mahdisoltani\textsuperscript{1} Mohammad Saniee Abadeh\textsuperscript{2}

\textsuperscript{1}Department of Computer Engineering, Sharif University of Technology, Tehran, Iran
\textsuperscript{2}Department of Electrical and Computer Engineering, Tarbiat Modares University, Tehran, Iran

Abstract

Image segmentation is one of the important components in many image-processing applications. Despite many researches on image segmentation, it still remains as an unsolved problem. This is mainly because identifying objects from an image data is hard task to do. In this paper, a new segmentation scheme is presented which segments the input color images in two soft and hard phases. After preprocessing the input image, a soft segmentation method is applied in order to segment the input image into initial small segments. Then, a hard segmentation phase starts by constructing a weighted network from the soft-segmented image and the communities of the network are extracted. Each resultant community in the hard segmentation phase represents a segment in the input image. Finally, a post-processing phase is done on the result of the hard segmentation phase. Parameter freeness is a very nice property that gives significance to HiSeg. To demonstrate the real pure result of HiSeg, a comprehensive sensitivity analysis is done on it. In addition, the results of HiSeg are demonstrated, and compared with some existing segmentation algorithms qualitatively and quantitatively. Extensive experiments have been performed and the results show that HiSeg can reliably segment the input color image into good subjective criteria.

Keywords: Image Segmentation, Color Image, Community Detection, Weighted Network.

1. Introduction

Color image segmentation is one of the most important problems in image processing. It is defined as the problem of partitioning pixels of an image into a set of homogeneous and meaningful regions such that the pixels in each region introduce same set of properties. There are many different techniques in the area of image segmentation, which can be broadly classified into histogram based, edge based, region based, clustering based, and combination of these techniques [1, 2]. There are many literatures on the color image segmentation some of which are reviewed to trace fresh study in the color image segmentation [3]. In [4] an efficient Lloyd’s k-means clustering algorithm requiring a kd-tree as the data structure is proposed. A pixon-based image segmentation model in combination with a Markov random field (MRF) model is proposed in [5]. Methods based on mixture of probability density functions defined in a multi dimensional image intensity space [6], Rough-set theory [7], fuzzy homogeneity and data fusion techniques [8], local kernel histograms in different illumination invariant color spaces [9] and Cellular Learning Automata (CLA) [10, 11] are some fresh studies in image segmentation.

In this paper we present a segmentation scheme in which acts in two soft and hard phases. After doing an arbitrary preprocessing step on input image, a soft segmentation method is applied to the preprocessed image to segment the input image into initial small segments. Afterwards, a hard segmentation phase starts by constructing a weighted network and extracting the communities of the network. Each resultant community in the hard segmentation phase represents a segment in the input image. Finally, an arbitrary post-processing is done on the result of the hard segmentation. Being parameter free is very nice property that gives merit to HiSeg. The experimental results show the superior segmentation capability of HiSeg over some well-known segmentation methods. The rest of the paper is organized as follows. Section 2 presents a review on related works on segmentation and Section 3 gives a brief overview.
of Community Detection. Section 4 presents the proposed algorithm for segmentation. Experimental results are presented in Section 5 comparing the proposed method against other methods. Section 6 gives the concluding remarks and future work.

2. Related Work

Segmentation is an important task in image and video processing that plays an important role in understanding images and videos. A variety of algorithms have been proposed in the literature for segmentation purposes. In [12] the existing methods in image segmentation are classified into three major categories: (1) feature-space-based clustering, (2) spatial segmentation, and (3) graph-based approaches. In feature-space-based clustering approaches, image features (usually based on color or texture) are used [13-16]. A specific distance measure is used to group the feature samples into compact, well-separated clusters ignoring the spatial information. Data clustering approaches are used in finding image features, but have some serious drawbacks as well. The spatial structure and the detailed edge information of an image are not preserved and pixels from disconnected regions of the image may be grouped together if their feature spaces overlap.

The spatial segmentation method is also referred to as region-based when it is based on region entities. Methods in this category gather similar pixels according to some homogeneity criteria [17, 18]. They are based on the assumption that pixels belonging to same homogeneous region, are more alike than pixels from different regions. The split-and-merge and region growing techniques are examples of such methods [19-21]. The watershed algorithm [22] is an extensively used algorithm for this purpose. However, it may undesirably produces a large number of small but quasi-homogenous regions, which demands some merging algorithm [16, 23].

Graph-based approaches are based on the combination of features and spatial information. In these approaches, grouping is based on factors such as similarity and continuation. The common idea among all these approaches is the formation of a weighted graph, where each vertex corresponds to an image pixel or a region, and the weight of each edge connecting two vertices represents the likelihood that they belong to the same segment. The constructed graph is partitioned into multiple components which minimize some cost function of the vertices in the components or the boundaries between them [24-28]. Other than the above-mentioned categories, hybrid approaches have emerged. Many of these hybrid techniques combine region-based methods with feature-based ones. These algorithms are popular for segmentation because they rely on both global and local information. The watershed algorithm [22] is an example of these hybrid algorithms. It begins by using a feature-based method to calculate the gradient magnitude and produces regions by a region-growing technique.

3. Louvain Community Detection Method

Community structure; which is a property of complex networks, can be described as the partitioning of a network into strongly connected groups such that there is a higher density of edges within groups than between them. Community structure detection has been one of the most popular research areas in recent years due to its applicability to a wide scale of disciplines. A network with community structure is shown in Figure 1. As this figure shows, there are three communities; denoted by the dashed circles, which have dense internal links but between which there are only a lower density of external links.

Several methods for community detection have been developed with varying levels of success. Minimum-cut method [29], Girvan-Newman algorithm [30], Modularity maximization [31] and Clique based methods [32]. An interested reader is referred to detailed surveys [33].

The Louvain community detection method [34] is a greedy optimization method, which is now one of the most widely used methods. The Louvain method consists of two phases: 1) Optimizing modularity in a local manner in order to find small communities. 2) Building a new network whose nodes are the communities.

The above steps are repeated iteratively until the maximum modularity is obtained. The Louvain method is applied as segment hierarchies extractor in our method, therefore it is described more detailed below. The algorithm starts with a weighted network of N nodes. First, each node of the network is assigned a different community. So, initially there are N communities in the network. Then, for each node i, the neighbour j of i is considered and the gain of modularity, which would take place by inserting i into the community of j is evaluated. The node i is placed in one of the communities of its neighbours for which the gain is maximum, but only if this gain is positive. If no positive gain is possible, node i stays in its original community. This process is applied repeatedly for all nodes until no further improvement is achieved. The first phase is then completed. The gain in modularity ΔQ obtained by moving i in the community C of j can easily be computed by:

$$\Delta Q = \frac{\Sigma_{in}^i + k_{i,in}}{m} - \left(\frac{\Sigma_{tot} + k_i}{2m}\right)^2 - \left(\frac{\Sigma_{in}^i}{2m} - \left(\frac{\Sigma_{tot}}{2m}\right)^2 - \left(\frac{k_i}{2m}\right)^2\right)$$

(1)

where $\Sigma_{in}$ is the sum of the weights of links inside $C$, $\Sigma_{tot}$ is the sum of the weights of links incident to nodes in $C$, $k_i$ is the sum of the weights of links incident to node $i$, $k_{i,in}$ is the
sum of the weights of the links from \( i \) to nodes in \( C \) and \( m \) is the sum of the weights of all links in the network. In the second phase of the algorithm, a new network is built whose nodes are the communities found in the first phase. The weight of a link between two new nodes is obtained by summing the weights of the links between nodes in the two corresponding communities. Then the first phase of the algorithm is reapplied to the resulting weighted network. The algorithm naturally incorporates the notion of hierarchy, as communities of communities are built during the process (see Figure 2). By construction, the number of meta-communities decreases at each time step, until there are no more changes and a local maximum is attained [36].

Figure 2. An illustrative example of the Louvain community detection algorithm. The colors show the first level partition and the surrounded clusters of nodes correspond to the partition at the second level [34]

4. Proposed Method: HiSeg

The proposed segmentation method is described in this section. Firstly, the input image is preprocessed in order to become ready for high quality segmentation. Afterwards, a soft segmentation method is applied to the preprocessed image and the image is segmented into initial small segments. Then the hard segmentation phase starts with constructing a weighted network and extracting the communities of the network. Each resultant community in the hard segmentation phase represents a segment in the input image. Finally, a post-processing is done on the result of hard segmentation phase. Figure 3 illustrates the structure of HiSeg. The details for each phase are described below.

4.1. Preprocessing

The preprocessing phase can be done if it is needed. This phase can include any common preprocessing algorithms such as noise removal, image filtering, image smoothing, etc.

4.2. Soft Segmentation

This phase include partitioning the input color image into possibly small regions, which are used in the hard segmentation phase to build the network. The soft segmentator should be adjusted to consider a trade-off between small and large size segments. Very small segments affect the time complexity of hard segmentation phase and large size segments lead the result of HiSeg to the result of soft segmentation phase. Any segmentation method, such as JSEG [35], super-pixel [36], meanshift [37], watershed [22] and levelset [38] can be used in this phase. In this paper, the meanshift is chosen as the soft segmentator. Figure 4 shows the result of this phase on a typical image.

Figure 4. (a) A typical color image, (b) The result of soft segmentation phase on (a)

4.3. Hard Segmentation

This phase contains two major steps. At first, a weighted network is constructed from the result of the previous step. Then the Louvain community detection algorithm is applied to the constructed network in order to detect the community hierarchy, which is then supposed as the segmentation result. Each step is described in more detail below. Two models can be used in the first step. As said above, a weighted network is constructed from the result of the soft segmentation phase. In the first model, each soft segment is considered as a node in the network, which is connected to its strict neighbours only. In the second one, each soft segment is considered as a node connected to all other existing nodes, which leads to complete graph. Figure 5 shows these two network models.

Figure 5. Two network construction models: (a) Strict neighbours model, (b) Complete graph model

The color histogram similarity is used to make the constructed graph weighted. For each pair of nodes connected through an edge, the similarity of them is calculated and assigned as the weight of the corresponding
edge. We use the RGB color space for the histograms. Each color channel is uniformly quantized into K levels and then the histogram of each region is calculated in the feature space of $K \times K \times K = K^3$ bins. The Jeffrey divergence or Jensen-Shannon divergence (JSD) is used to compare histograms:

$$d_{JSD}(H, H') = \sum_{m=1}^{M} \frac{H_m \log \frac{2H_m}{H_m + H'_m}}{H_m + H'_m} + \frac{H'_m \log \frac{2H'_m}{H_m + H'_m}}{H_m + H'_m}$$

where $H$ and $H'$ are the histograms to be compared and $H_m$ is the mth bin of $H$. The value of $1 - d_{JSD}$ is considered as the similarity value which is assigned to the edge connecting two regions.

After constructing the weighted network, the Louvain community detection algorithm detects a hierarchical community structure of the network, as mentioned in section 3. When the Louvain community detection algorithm meets its stopping condition, the network is grouped into different communities in two hierarchies. Since each node in the network represents a soft segment, each extracted community represents a more general segment as a combination of some soft segments. Actually, the Louvain community detection method does a hard segmentation on the soft-segmented image in a network structure. Figure 6 shows the result of soft segmentation along with two hierarchy levels of hard segmentation phase on a typical image. As this figure shows, the result of second level is more precious than the first one.

Figure 6. (a) Input image, (b) Soft segmented image, (c) First level hard segmented image, (d) second level hard segmented image

Being parameter freeness of the Louvain community detection algorithm is a nice property, which gives significance to HiSeg. This property makes HiSeg able to segment the input images with no need to input parameters.

4.4. Post Processing

This phase is applied to the result of the hard segmentation phase, and merges small segments with the most similar and probable neighbour ones in order to enhance the subjective quality of the result. This phase is arbitrary and can be ignored along the segmentation process.

5. Experimental Results

We have performed our experiments successfully using HiSeg algorithm on images selected from Berkeley segmentation dataset (BSDS) [39]. The method is carried out on a 2 GHz processor with 1024 MB RAM on Windows XP professional platform. MATLAB 7.1 and image processing toolbox 5.0.2 are used. No pre and post processing are done to demonstrate the pure result of HiSeg. Anyway, these steps are arbitrary and can be applied when the input image requires preprocessing. The performance analysis of HiSeg is done in two sensitivity analysis and comparison steps. Each step is described in more details below.

5.1. Sensitivity Analysis

The sensitivity analysis is done to demonstrate the robustness and subjective quality of the algorithm to different model parameters. The quantization levels K is set to 16 during all the analysis process. The sensitivity is considered with two network modelling paradigms (strict neighbours and complete graph), two different values as the minimum size allowed of the regions resulted from soft segmentation phase (750 and 3000), and two different community detection algorithms in the hard segmentation phase (the Louvain community detection algorithm and Affinity Propagation (AP) clustering method [40]). Also some more sensitivity analysis can be done on the effect of noise and other filters.

Figure 7 shows the results of HiSeg on some typical images with the two mentioned different network modelling paradigms. The quantization levels K is set to 16 and the soft segmentation is done with minimum allowed size of 750. The Louvain community detection algorithm is used as the hard segmentation algorithm. The results have approved the robustness and accuracy of the complete network model in clustering process. So, the second model (The complete model network) is selected and applied to model the network and to analyze the sensitivity of HiSeg to other parameters. Table 1 shows the configuration for analysing the sensitivity of HiSeg to different network modelling paradigms.

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantization levels</td>
<td>16</td>
</tr>
<tr>
<td>Minimum size allowed</td>
<td>750</td>
</tr>
<tr>
<td>Community detection algorithm</td>
<td>Louvain algorithm</td>
</tr>
<tr>
<td>Preprocessing</td>
<td>No</td>
</tr>
<tr>
<td>Post-processing</td>
<td>No</td>
</tr>
</tbody>
</table>

Table 1. Configuration for sensitivity analysis to different network modelling paradigms

To analyze the sensitivity of HiSeg to the minimum size allowed in the soft segmentation phase, two values of this parameter are considered, 750 and 3000. The result is presented in Figure 8. The quantization levels K is set to 16 as before, and the complete network model is used. In addition, the Louvain community detection algorithm is used in the hard segmentation phase. As the results show, HiSeg does not have enough freedom for analysing the constructed network and detecting the community hierarchies when the initial region size are large. So, the minimum size allowed 750 is selected and applied to model the network and analyze the sensitivity of HiSeg to the other parameters.
Table 2 shows the configuration for analysing the sensitivity of HiSeg to two different minimum sizes allowed.

Table 2. Configuration for sensitivity analysing to two different minimum sizes allowed

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantization levels</td>
<td>16</td>
</tr>
<tr>
<td>Network modeling</td>
<td>Complete model</td>
</tr>
<tr>
<td>Community detection algorithm</td>
<td>Louvain algorithm</td>
</tr>
<tr>
<td>Preprocessing</td>
<td>No</td>
</tr>
<tr>
<td>Post-processing</td>
<td>No</td>
</tr>
</tbody>
</table>

To show the superiority of the Louvain community detection method, the subjective quality of HiSeg is demonstrated by replacing Louvain community detection method with Affinity Propagation (AP) clustering method [40]. AP defines an energy-based formulation of the K-centers combinatorial optimization problem, which is solved using a message passing algorithm akin belief propagation. The AP starts with a distance matrix of inputs and result a number of clusters, which minimize the energy function. The quantization levels $K$ and the minimum size allowed in soft segmentation phase are set to 16 and 750 respectively.

The complete network model is selected as well. Figure 9 shows the result of HiSeg by replacing Louvain community detection method with the Affinity Propagation (AP) clustering method. As the figure shows, the Louvain community detection method demonstrates more qualitative result than the Affinity Propagation method. Table 3 shows the configuration for analysing the sensitivity of HiSeg to two Louvain community detection and Affinity Propagation (AP) clustering method.

Table 3. Configuration for sensitivity analysing to two Louvain community detection and Affinity Propagation (AP) clustering method

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantization levels</td>
<td>16</td>
</tr>
<tr>
<td>Minimum size allowed</td>
<td>750</td>
</tr>
<tr>
<td>Network modeling</td>
<td>Complete model</td>
</tr>
<tr>
<td>Preprocessing</td>
<td>No</td>
</tr>
<tr>
<td>Post-processing</td>
<td>No</td>
</tr>
</tbody>
</table>

5.2. Comparison

In this section the results of HiSeg are demonstrated and compared with the results of some existing segmentation algorithms. By the result acquired from sensitivity analysis section, the parameters of HiSeg are adjusted. The quantization levels $K$ and the minimum size allowed in soft segmentation phase are set to 16 and 750 respectively. The complete network model is selected and the Louvain community detection method is used as the hard segmentator. Figure 10 shows the result of HiSeg on some images selected from Berkeley segmentation dataset (BSDS) based on configurations given in table 4.

Figure 7. Comparison of two network modelling paradigms on segmentation result: (a) Soft segmented image, (b) Segmentation result on strict neighbours model network, (c) Filled segmentation result on strict neighbours model network, (d) Segmentations result on complete model network, (e) Filled segmentation result on complete model network
To compare HiSeg with other methods, the results are compared with three segmentators JSEG [35], EDISON [41] and MULTISCALE [42]. These segmentation methods are well-known and often used for image segmentation. The JSEG [35] method separates the segmentation process into two independently processed stages: color quantization and spatial segmentation. In the first stage, colors in the image are quantized to several representing classes that can be used to differentiate regions in the image. A region growing method is then used to segment the image based on the multi-scale property. The JSEG is able to establish its parameters automatically. Therefore, we did not set essential parameters at all while testing JSEG.

EDISON program system [41] implements the mean shift segmentation algorithm [37] in both boundaries extraction and noise filtering scheme. One of the main parameters of EDISON is the minimal region size (in pixels) which this method can create. For testing, we set values of this parameter to 1000. Other parameters have the same values that the authors used as default [41]. In MULTISCALE segmentation method [42], an image first is analyzed in coarser scale, and then in finer scale. In [42], some parameters are recommended as "safe". These values have been taken at testing. Figure 11 compares the results of HiSeg with JSEG, EDISON and MULTISCALE qualitatively, which indicate the superiority of HiSeg over other methods.

For quantitative evaluation, we investigate the widely used Probabilistic Rand Index (PRI) [43]. The PRI measures the consistency of labellings between a segmentation and its ground truth by the ratio of pairs of pixels having the same labels, averaging across multiple ground truth segmentations to account for variation in human perception. This measure takes the values in the interval [0, 1]; more is better. Table 5 depicts the mean of the PRI values that are calculated when the JSEG, EDISON, MULTISCALE and proposed method were applied to all 300 images in the Berkeley segmentation dataset (BSDS).

As stated before, HiSeg uses the color histogram as a measure of similarity. Therefore, it is unlikely for the approach to operate properly where color similarity is high between the regions with different textures (See Figure 12). Anyway, many segmentation methods often encounter problem in such cases and try to use the texture information between regions to overcome this problem. As figure 12 shows, HiSeg failed to segment the desired object properly. The high similarity between the object and the background in color image causes such poor results. Hence, incorporating the texture information can be considered as a good candidate to handle such cases.

6. Conclusions

The methodology presented in this paper aimed in improving image segmentation based on initial soft segmentation of regions and doing a hard segmentation phase consequently to extract the final segments. The soft segmentation phase involves segmentation of the input image into small regions by means of meanshift segmentation algorithm. The hard segmentation phase consists of constructing a weighted network on the soft-segmented image and extracting the community hierarchies of the network. Each resultant community in the hard segmentation phase represents a segment in the input image.

Through using a scheme for unfolding different clusters, HiSeg could utilize the power of community detection paradigm to generate robust clustering results and therefore achieves robust image segmentation. Experimental results show that HiSeg works better than some well-known image segmentation methods in both subjective and quantitative criterions. A drawback of HiSeg is that, it does not consider the texture information. So, it fails to make good results when adjacent regions with similar color histograms are different in texture. In our future work, we will incorporate also some texture information to improve HiSeg.

References


Figure 8. Comparison of different minimum sizes allowed in the soft segmentation on segmentation result: (a) Input image, (b) Soft segmentation with minimum size allowed 750, (c) Filled final result on (b), (d) Soft segmentation with minimum size allowed 3000, (e) Filled final result on (d)
Figure 9. Comparison of two Louvain community detection method and Affinity Propagation (AP) clustering method on segmentation result: (a) Input image, (b) Soft segmented image, (c) Segmentation result by applying the Affinity Propagation method on (b), (d) Segmentation result by applying the Louvain community detection method on (b), (e) Filled segmentation result by applying the Louvain community detection method on (b)
Figure 10. The result of HiSeg on some images selected from Berkeley segmentation dataset (BSDS) based on configurations given in table 4: (a) Input image, (b) Soft segmented image, (c) Final segmentation result, (d) Filled final segmentation result
Figure 11. Qualitative comparison of the results of HiSeg with JSEG, EDISON and MULTISCALE. (a) Input image, (b) JSEG [35], (c) EDISON [41], (d) MULTISCALE [42], (e) HiSeg
Figure 12. Investigating the behaviour of HiSeg in the presence of high color similarity among regions with different textures (a) Input image, (b) Soft segmented image, (c) Final segmentation result, (d) Filled final segmentation result


Ahmad Ali Abin received a B.Sc. in computer engineering from Iran University of Science & Technology, Iran, in 2005. In September 2008, he completed the M.Sc. degree in computer engineering at Sharif University of Technology, Iran. Currently he is a Ph.D. student at the Department of Computer Engineering, Sharif University of Technology. His research interests focus on pattern recognition, machine learning, neural computing and image processing.

E-mail: abin@ce.sharif.edu

Farzaneh Mahdisoltani received her B.Sc. and M.Sc. degrees from the Department of Computer Engineering, Sharif University of Technology, in 2010 and 2012 respectively. She is currently a visiting scholar at Max Planck Institute. She worked as a research intern at Ecole Polytechnique Federale de Lausanne (EPFL) in 2012. Her research interests include machine learning, and data mining.

E-mail: mahdisoltani@ce.sharif.edu

Mohammad Saniee Abadeh received his B.S. degree in Computer Engineering from Isfahan University of Technology, Isfahan, Iran, in 2001, the M.S. degree in Artificial Intelligence from Iran University of Science and Technology, Tehran, Iran, in 2003 and his Ph.D. degree in Artificial Intelligence at the Department of Computer Engineering in Sharif University of Technology, Tehran, Iran in February 2008. Dr. Saniee Abadeh is currently a faculty member at the Faculty of Electrical and Computer Engineering at Tarbiat Modares University. His research has focused on developing advanced meta-heuristic algorithms for data mining and knowledge discovery purposes. His interests include data mining, bio-inspired computing, computational intelligence, evolutionary algorithms, fuzzy genetic systems and memetic algorithms.

E-mail: saniee@modares.ac.ir

Paper Handling Data:
Submitted: 16.07.2011
Received in revised form: 24.05.2013
Accepted: 17.07.2013
Corresponding author: Ahmad Ali Abin,
Department of Computer Engineering, Sharif University of Technology, Tehran, Iran.