Split-and-Merge Segmentation Employing Thresholding Technique

Hee Soo Yang and Sang Uk Lee

Real Time Vision Laboratory, School of Electrical Engineering
Seoul National University, Seoul, 151-742, KOREA
E-mail {hsyang, sanguk}@sting.snu.ac.kr

Abstract

The conventional split-and-merge algorithm is lacking in the adaptability to the image semantics because of its stiff quadtree-based structure. In this paper, a thresholding technique is employed in the splitting phase of the split-and-merge segmentation scheme to directly reflect the image semantics to the image segmentation results. Thus, the regions which contain distinct subregions can be extracted by one step, which significantly alleviating the computation load and the memory requirements. To overcome the problems, caused by the block-based thresholding, such as edge discontinuities over the borders of quadtree nodes, a prefilter is used to preserve and sharpen the edge information. Applying the proposed algorithm to aerial images results in the reduction of the number of intermediate regions generated in the splitting phase, demonstrating the improved ability to capture the image semantics. The simulation results show that the proposed algorithm successfully extract arbitrary shaped regions in aerial images.

1 Introduction

Quadtree-structured split-and-merge is a popular approach for image segmentation because of its simplicity and computational efficiency. An inherent drawback related to the quadtree-structured split-and-merge segmentation is its inability to adapt to the image semantics. The edges of the segments, formed by this algorithm, can have only two orientations (horizontal and vertical), and their positioning is restricted by the borders of the quadtree nodes. It is difficult to achieve the agreement between the rigid rectilinear segment boundaries and image edges of arbitrary orientation or position. These errors in edge orientation and position inevitably yield errors in segmentation.

Wu proposed four-way split algorithm to adapt to the image semantics[2]. In Wu's approach, bipartitions by horizontal, 45° diagonal, vertical, and 135° diagonal cuts are considered to improve segmentation validity over regular rectilinear cuts. But this approach cannot reproduce the edges with arbitrary orientation.

In this paper, we propose a modified split-and-merge algorithm to extract regions from aerial images. The modified algorithm employs thresholding technique to extract regions whose boundaries are arbitrary oriented, shaped and positioned. The proposed algorithm is shown to be very efficient in reducing the number of iterations in the split process, and the number of regions generated by the split process, requiring considerably less memory and computational complexity.

2 Segmentation algorithm

In the first phase of the conventional SM (split-and-merge), the image is initially considered as a single region, and if considered non-homogeneous by a dynamic range criterion, then it is split into four regions (according to a quadtree structure). The splitting algorithm is recursively applied to each of the resulting regions, until the homogeneous criterion is fulfilled. The second phase of the SM is region merging, in which pairs of adjacent regions are compared and merged, if their union satisfies the homogeneity criterion.

We approximate each region to a constant planar surface, and use the mean squared error measure, given by

\[ d(S_i) = \frac{1}{|S_i|} \sum_{(x,y) \in S_i} (I(x,y) - \mu(S_i))^2 \]  (1)

where \( I(x,y) \) is the intensity of the pixel \((x,y)\), \( S_i \) is a connected subset of the image, and \( \mu(S_i) = \sum_{(x,y) \in S_i} f(x,y)/|S_i| \) is the average intensity of the segment \( S_i \), respectively. If the mean squared error \( d(S_i) \) is below (above) a threshold \( \epsilon > 0 \), then the corresponding region is discriminated to be homogeneous (non-homogeneous).

In this paper, the region splitting strategy is modified, to reproduce arbitrary shaped region in one step.
As mentioned previously, in the conventional scheme, a given region is recursively split into four subregions, until the region is considered homogeneous. Recursive split will fulfill the homogeneous condition, whenever the region size reaches one. Therefore, in the splitting phase of the conventional SM, a region is considered either homogeneous where no more splitting process is needed, or nonhomogeneous where the given region should be split into subregions. There are three types of regions in the modified splitting algorithm. We first check the homogeneity of the given region. If the region satisfies homogeneous criterion, then the region is considered as a homogeneous region. For each non-homogeneous region, we apply a thresholding technique to the region, yielding the thresholded subregions. Then the subregions generated by thresholding are examined their uniformity. If all of the subregions are homogeneous, then the given region is labeled as a thresholded region, and no more splitting process will be applied to that region. If at least one of the subregions is nonhomogeneous, then the whole splitting process is applied to the region again, recursively. The blockdiagram for the modified split algorithm is shown in Figure 1.

Among various thresholding techniques, in our approach, Otsu’s method [1] is used to select the threshold. The threshold is selected by the discriminant criteria, which attempt to maximize the separability or minimize the within-class variance, given by

$$\sigma_w^2 = \frac{|S_1|}{|S_1| + |S_2|} \sigma_1^2 + \frac{|S_2|}{|S_1| + |S_2|} \sigma_2^2$$  \hspace{1cm} (2)

where $\sigma_1^2$ and $\sigma_2^2$ are the variances of the resultant segment 1 and 2, respectively. And $\sigma_w^2$ is a weighted sum of the variances of the two regions which are generated by applying the threshold to the image. Since, Otsu’s method minimizes $\sigma_w^2$ and consequently $\sigma_1^2$ and $\sigma_2^2$, which are mean squared errors of the constant planar approximation, it is pertinent to our approach, since the mean squared error is used as a homogeneous measure.

As the thresholded regions are likely to contain more than two subregions, they need a labeling phase, in which each connected subregion is labeled as distinct regions. The regions which are not labeled as homogeneous nor thresholded are split into four subregions, and the process of splitting is recursively applied to each of the resulting subregions.

After the modified splitting process, we obtain nonhomogeneous regions and homogeneous regions. The nonhomogeneous regions are small and rectilinear blocks, which will not be produced, if single-pixel regions are allowed. There are two types of homogeneous regions; one is rectilinear-shaped block, generated by a splitting process based on quadtree structure, and the other is arbitrary-shaped region, created from a thresholding process.

To obtain the final segmentation, we need to analyze and merge adjacent regions, if their union satisfies the homogeneity condition. Two regions created by applying thresholding technique to a father region cannot be merged because otherwise, the region will not be split in the first place. This means that any two candidate regions to be merged should be adjacent across borders of the quadtree nodes.

In our approach, we adopt a greedy algorithm [2] to merge the regions generated by the splitting process, which always merges two adjacent regions whose merge incurs the minimum increase in the approximation error among all possible merges. A list of the region adjacency is constructed and manipulated to increase the efficiency of the merging phase. The two mergeable nodes of the lowest cost is combined into one node, corresponding to a new larger uniform region, and the list is updated accordingly. The terminating condition of the merging phase could be the number of the overall resultant regions, the approximation error of the merging regions, etc. We omit the detail of the terminating condition here, because our main work focuses on the modification of the split process.

Our approach is mainly based on the basic split-and-merge strategy, but its novelty is the ability to threshold a region of bimodal histogram into corresponding subregions.

3 Results and discussion

Before applying the proposed algorithm to aerial images, we first perform prefiltering to smooth the input image. The purpose of prefiltering is to reduce the noise level of the input image and to sharpen the edge information. Sharpening the edge information is crucial in our approach, in which thresholding is per-
formed in each quadtree nodes. If the slope of a edge is small, a slight change in the threshold yields a large variation of the region boundaries, which are likely to lead to edge discontinuities over the borders of the quadtree nodes. The symmetric nearest neighborhood filter is employed in our approach, which preserves and even more sharpens the edge information[4]. The pre-filtering is found to be very effective in extracting the smooth boundaries of segments over the block boundary.

The proposed segmentation algorithm is implemented in C and tested on various aerial images. The results of experiments shows that the proposed algorithm yields significant improvement in segmentation validity over the regular decomposition. In this paper, the quadtree-based regular decomposition algorithm [5] is chosen for comparison with the proposed algorithm. This algorithm recursively splits the input image into four equally sized rectangles.

In the experiments, the same mean-square error threshold is used for both algorithms to terminate the splitting of an image region. After the recursive split is completed, the same greedy merge is applied to both image partitions, generated by proposed and the regular split methods. Therefore, the two tested algorithms only differs in the split phase. We compare the final segmentation results by the two algorithms for the same number of segments, since the subjective criterion for image understanding is to satisfactorily characterize the image by the least possible number of segments.

Figure 2 shows the original aerial image and two segmentation results by the proposed and the regular decomposition algorithms, respectively. There are 40 segments in each of the two segmented images. The superiority of the proposed algorithm is clearly demonstrated by the results shown in Figure 2 (c). Notice the false segments a and d, which are merged, and segment d is not properly merged with the adjacent segment. Segment b and c are also falsely merged, while the proposed algorithm yields a correct result.

The adaptability of the proposed algorithm to the image semantics creates considerably fewer number of regions than the regular decomposition methods for the same mean-square error threshold. The number of generated split regions are provided in Table I. The proposed algorithm generates only about 60% of the split regions created by the conventional algorithm, alleviating the computational load of the merge step and hence of the overall algorithm[3].

Figure 3 shows the original image and two intermediate results obtained by the proposed and the regular splitting algorithms. The conventional quadtree-based algorithm yields 9858 regions, while the proposed algorithm yields 6098 regions of larger size, indicating that the proposed algorithm reproduces the image semantics successfully.

### Table 1: The number of regions generated by split process

<table>
<thead>
<tr>
<th></th>
<th>proposed</th>
<th>conventional</th>
</tr>
</thead>
<tbody>
<tr>
<td>image-1</td>
<td>3001</td>
<td>5661</td>
</tr>
<tr>
<td>image-2</td>
<td>4347</td>
<td>7053</td>
</tr>
<tr>
<td>image-3</td>
<td>6098</td>
<td>9858</td>
</tr>
</tbody>
</table>

4 Conclusion

A modified split-and-merge employing thresholding technique was proposed to improve the validity of the conventional split-and-merge segmentation algorithm. The new algorithm extracts regions whose boundaries are arbitrary oriented, shaped, and positioned. The proposed algorithm is based on the conventional split-and-merge strategy, but its novelty is the ability to threshold a region of bimodal histogram into corresponding subregions in one iteration step. The simulation results showed that the proposed algorithm successfully captures the image semantics in aerial images. The proposed algorithm extracts smaller number of regions whose size is bigger than the conventional quadtree-based algorithm, significantly alleviating the reduction of the computational load and the memory requirements.

### References


Figure 2: original image and segmented images

Figure 3: original image and split images