Proceedings of the Fifth IIEEJ International Workshop
on Image Electronics and Visual Computing 2017
Da Nang, Vietnam, February 28- March 3, 2017

FAST SUPER-PIXEL SEGMENTATION WITH CLEAR DETAIL AND SHORT BOUNDARY

Masayuki Miyama†
†Kanazawa University

ABSTRACT

Super-pixel segmentation is to over-segment an image to get object boundaries. This is placed for a preprocessing of the image segmentation and the object recognition. This paper proposes a novel fast and accurate super-pixel segmentation algorithm. The proposed method departs from the initial division of a regular grid and hierarchically repeats boundary update using an energy function in local. The energy function consists of a penalty term and a regularization term. The penalty term expresses the color similarity between a pixel and a super-pixel. The regularization term expresses the spatial continuity of the super-pixel. Experimental results show the processing speed is 2-10 times faster than that of the conventional methods without severe degradation in accuracy. The accuracy against the boundary length is overwhelmingly higher in particular. The object inside is meaningfully divided and the object detail is clearly shown in the image substituting the super-pixel average for each pixel. It is expected that these characteristics open up possibility to new applications such as parts recognition in the object or compact image expression.

1. INTRODUCTION

A super-pixel is a small set of adjacent pixels with similar characteristics such as colors. Super-pixel segmentation is to over-segment an image to get object boundaries. It is important that the super-pixel border accords with the object border, and the object inside may be over-segmented then. The super-pixel segmentation is a preprocessing of the image segmentation and the object recognition[1]. Execution time of the post-processing can be drastically reduced by adopting a super-pixel instead of a pixel as a processing unit.

The super-pixel segmentation is aimed for the accurate extraction of the object boundary. Little one for the number of partitions is desirable then. Because it is a preprocessing of the applications, the processing time should be short. There are two main approaches for the super-pixel segmentation. The graph approach expresses an image by a graph, and optimizes the graph-cut problem in global[2, 3]. This approach is generally high in precision, but low in speed. Another approach leaves from the initial division and repeats clustering in local[4, 5]. The speed of this approach is high, but the precision is low in general. Fast and accurate algorithms have been continuously developed until now.

In this paper, we propose a fast and accurate algorithm of super-pixel segmentation adopting the local approach. The experimental results show that the processing speed is 2 to 10 times faster than the conventional methods, and there is no remarkable decrease in accuracy. The accuracy with respect to the boundary length is overwhelmingly high. The details of the object are clearly shown in the image where the super-pixel average has been assigned to each pixel. This fact indicates that the inside of the object is divided according to its meaning.

This paper is organized as follows. The next section describes the conventional algorithms. The section three explains the proposed method. The section four describes the experimental setup and results. The section five concludes this paper.

2. CONVENTIONAL ALGORITHMS

There are two main approaches for the super-pixel segmentation. The first approach expresses an image by a graph and optimizes the graph-cut problem globally. This is accurate but slow in general. The second approach starts from the initial division and repeats the local optimization. This is generally inaccurate but fast.

ERS (Entropy Rate Super-pixel Segmentation) is a segmentation algorithm using a graph[3]. A node of the graph is a pixel, and an edge is a similar degree between two adjacent pixels. The problem is translated to divide the graph into subgraphs separated each other. The algorithm performs a random walk in the solution space and finds the most suitable solution according to an objective function. The function consists of two terms. The first term to express entropy rate is obtained by the uniformity of the pixels in the super-pixel. The second term is to balance each super-pixel size. This method is highly precise and high-speed for the global approach, but is slower than the local methods.

SLIC (Simple Linear Iterative Clustering) is a division algorithm using the k-means method[4]. When assigning a pixel to a super-pixel, the algorithm considers a distance from the pixel to the center of the super-pixel. Therefore, SLIC is characterized by the uniform size and shape of super-pixels. SLIC is high-speed but the precision is not so high.

SEEDS (Super-pixels Extracted via Energy-Driven Sampling) also adopts the local approach[5]. The algorithm generates an image pyramid, then divides the most upper image into the objective number of partitions in a lattice form. The boundary update is repeated on every level of hierarchy from the top to the bottom. The color histogram of the super-pixel is compared with that of the pixel on the division border. The pixel is assigned to the most similar super-pixel. It is relatively high in the precision and faster than ERS, but slower than SLIC.

3. PROPOSED ALGORITHM

The proposed method is based on SEEDS. Similar to SEEDS, the proposed method departs from the initial division of a regular grid, and repeats boundary update hierarchically as shown in Fig.1a-1c. SEEDS has high computational complexity and large memory consumption to calculate the similarity between a pixel and a super-pixel due to the adoption of the color histogram. The proposed method simply uses a pixel value instead of the histogram to overcome the disadvantages of SEEDS. We call our method as simplified SEEDS (SS) later in this paper.

SS scans an image in raster order and updates boundaries on each level of hierarchy from the top to the bottom. The target pixel on the boundary is compared to the super-pixels appeared in the 3 by 3 patch of the target. The following energy function is used for the update.

$$E(p,s) = E_p(p,s) + \alpha E_r(s) \tag{1}$$

Here, p denotes a pixel, and s denotes a super-pixel. The energy function consists of a penalty term E_p and a regularization term E_r. α is a coefficient to balance between E_p and E_r. The penalty term E_p is written as follows.

$$E_p(p,s) = \sum_{c \in C} |I_c(p) - I_c(s)| \tag{2}$$

c is a color component of a color system C, and $I_c(p)$ means the component intensity of the pixel. This term expresses the similarity between the target pixel and the super-pixel. This term becomes small as their difference is small. The regularization term is written as follows.

$$E_r(s) = 8 - \sum_{p \in P} \delta(L(p), s) \tag{3}$$

p is a pixel in a 3 by 3 patch P, $L(p)$ returns the super-pixel to which p belongs, and the function $\delta(a,b)$ returns 1 when a equals to b otherwise 0. This term expresses the spatial continuity of the super-pixel including the target pixel. This term becomes small as the number of pixels belonging to the super-pixel is large. When updating boundaries, the target pixel is assigned to the super-pixel with the smallest energy using this function, as shown in Fig.1d.

4. EXPERIMENTS

4.1. SETUP

We compared our SS with SEEDS, SLIC, and ERS. The open source programs of the conventional algorithms developed by the original authors were used. The parameter values for each algorithm were left as originals. The program of SS was developed on Miscrosoft VisualStudio 2013. The coefficient α in the energy function was eight. The number of iterations to scan an image in raster order for boundary update on each hierarchical level was two. The Berkeley segmentation benchmark, including 200 test images, was used in the experiment[6]. The Matlab toolbox developed by [1] was used for automatic benchmarking and evaluation. The programs run on a PC with Intel Core i5 CPU at 3 GHz and 4 GB memory.

We adopted the standard metrics of boundary recall (BR), and undersegmentation error (UE) to evaluate the algorithms. The BR is a recall rate of the object boundary. It is expressed as follows.

$$BR = \frac{TP}{TP + FN} \tag{4}$$

TP stands for the true positive that means the number of boundary pixels determined by the algorithm consistent with the ground truth. FN stands for the false negative that means the number of non-boundary pixels determined by the algorithm inconsistent with the ground truth. A high BR value is preferred.

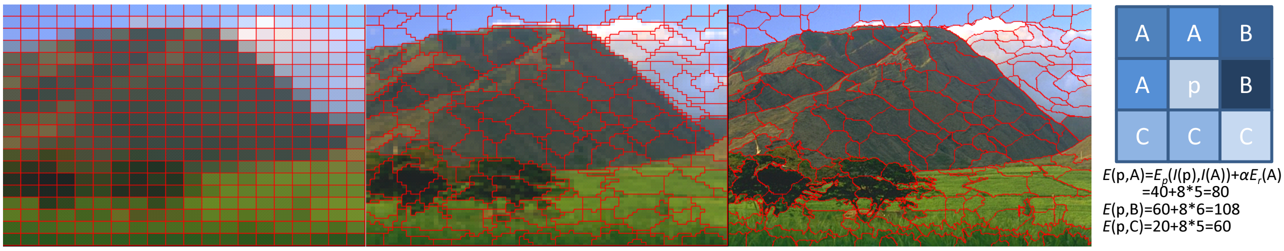
The UE is a ratio of the total area overlapping the object boundary to the whole image. It is expressed as follows.

$$UE = \frac{1}{N}\left[\sum_{S \in GT}\left(\sum_{P:P\cap S \neq \emptyset} \min(P_{in}, P_{out})\right)\right] \tag{5}$$

N is the number of pixels in an image, S is a segment in the ground truth GT, and P is a super-pixel produced by the algorithm. P_{in} is the area of P inside of S. P_{out} is the area of P outside of S. A low UE value is preferred.

4.2. RESULTS

Fig.2 shows the experimental results for the 200 test images in the benchmark. Fig.2a-2b show the relation between the number of segments(NoS) and accuracy metrics. SS was more accurate than SLIC. It was less accurate than ERS



(a) Level 4 (b) Level 2 (c) Level 0 (d) Energy function

Figure 1: Proposed algorithm

and SEEDS, but the degradation was small. Fig.2c-2d show the relation between the boundary length(BL) and accuracy metrics. The segmentation accuracy against the boundary length of SS was overwhelmingly higher than that of others. This is because the proposed method simply divides an image with the color similarity. Our method does not consider the uniformity in area and shape of the super-pixel. A super-pixel overlapping the object boundary is eroded by adjacent super-pixels existing on the object inside, and finally disappears. Then useless borders disappear, and object borders are left. The accuracy against the boundary length of our algorithm becomes higher in this way. Fig.2e shows the relation between NoS and the runtime(RT). SS was faster than ERS by 10 times, SEEDS by 3 times, and SLIC by 2 times.

Fig.3 shows result images for a test of No.145059. The images on the left side were obtained by replacing each pixel value with the average value of the corresponding super-pixel. Segmentation boundaries are appended to the right images. Face parts, the striped pattern of wears, and the pattern of the sword handle were clearly shown by SS in Fig.3. In spite of about the same number of segments as the other methods, SS generated an image with clear details and short boundaries. Similar results were obtained with other test images.

5. CONCLUSION

This paper proposed a novel fast and accurate super-pixel segmentation algorithm. The proposed method starts from the initial division of the regular grid and repeats the boundary update hierarchically using the local energy function. The energy function consists of a penalty term to express the color similarity between a pixel and a super-pixel and a regularization term to express the spatial continuity of the super-pixel. Experimental results showed the proposed method was 10 times faster than ERS without severe degradation in accuracy. It was 2 times faster and more accurate than SLIC. The accuracy with respect to the boundary length was overwhelmingly high. Since the inside of the object is divided according to the meaning, the details of the object in the image where each pixel was replaced with a super-pixel average were obvious. These features are expected to open up possibilities for new applications such as component recognition within objects and compact image representation.

ACKNOWLEDGMENT

This study received the furtherance of Grant in Aid for Scientific Research (15K00227).

[1] P. Neubert and P. Protzel: “Superpixel Benchmark and Comparison,” In Proc. Forum Bildverarbeitung (2012).

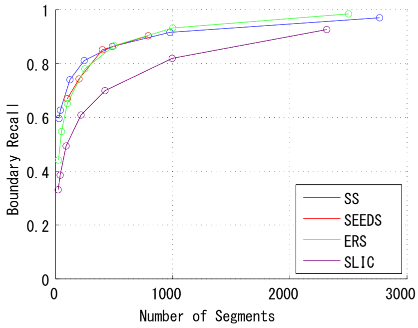
[2] PEDRO F.FELZENSZWALB and DANIEL P.HUTTENLOCHER: “Efficient Graph-Based Image Segmentation,” International journal of Computer Vision 59(2), 167-181 (2004).

[3] M.-Y.Liu, O.Tuzell, S.Ramalingam, and R.Chellappa: “Entropy Rate Superpixel Segmentation,” IEEE Computer Vision and Pattern Recognition (CVPR) (2011).

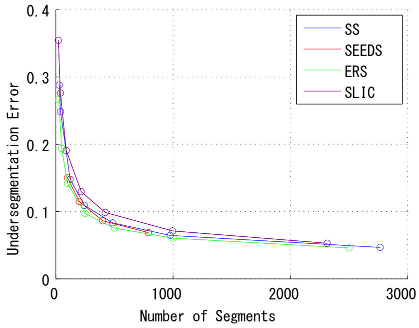
[4] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Susstrunk: “SLIC Superpixels,” EPFL Technical Report no. 149300, June (2010).

[5] Michael Van den Bergh, Xavier Boix, Gemma Roig, Benjamin de Capitani, and Luc Van Gool: “SEEDS: Superpixels Extracted via Energy-Driven Sampling,” ECCV 2012, Part VII, LNCS 7578, pp.13-26 (2012).

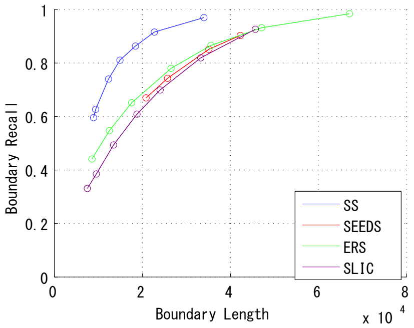
[6] D.Martin, C.Fowlkes, D.Tal, and J.Malik: “A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics,” Proceedings. Eighth IEEE International Conference on Computer Vision (Volume:2) (2001).



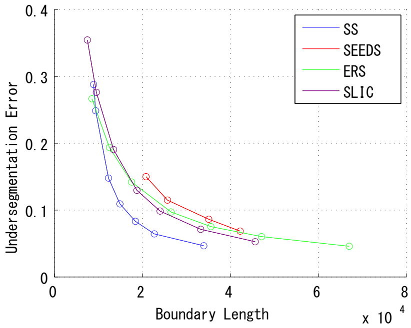
(a) NoS vs. BR



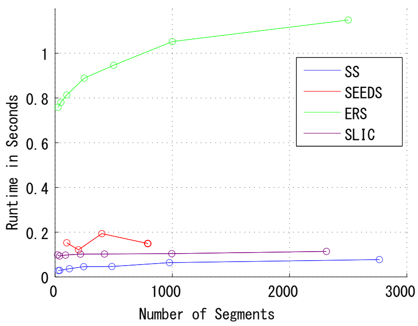
(b) NoS vs. UE



(c) BL vs. BR



(d) BL vs. UE



(e) NoS vs. RT

Figure 2: Experimental results

(a) SS (NoS:915, BL:24964)

(b) ERS (NoS:1000, BL:49099)

(c) SLIC (NoS:1052, BL:35658)

(d) SEEDS (NoS:796, BL:41178)

Figure 3: Result images of test No.145059