TOWARDS ROBUST AND EFFICIENT SEGMENTATION: AN APPROACH BASED ON INTER-REGION CONTOUR AND INTRA-REGION CONTENT ANALYSIS

Zhiding Yu, Oscar C. Au, Ketan Tang, Lingfeng Xu, Wenxiu Sun and Yuanfang Guo

Hong Kong University of Science and Technology
Department of Electronic and Computer Engineering
Clear Water Bay, Kowloon, Hong Kong
Email: {zdyu, eeau, tkt, lingfengxu, eeshine, eeandyguo}@ust.hk

ABSTRACT

We address the problem of boundary estimation by formulating it as inter-region contour and intra-region information analysis in the framework of graph-based segmentation. Given an image without any prior information about object model and class, we seek to approximate one's instant perception of visual similarity. The method can serve as a preprocessing step for many higher level operations that require regional support, such as scene understanding and object recognition. We show in this paper that the defined region comparison predicate makes a better boundary estimator than efficient graph-based image segmentation (EGS) - a well known and widely used segmentation method. We further illustrate, by making a small relaxation, further improvement of segmentation performance can be achieved. Experimental results have demonstrated the effectiveness of our proposed method.

***Index Terms*—** Image Segmentation, Graph-based method, Boundary Estimation

1. INTRODUCTION

In computer vision, the problem of segmentation and perceptual grouping remains challenging despite years of extensive study. The goal of segmentation is to assign the same labels to spatially connected pixels that share similar characteristics with certain predefined criteria. From the perspective of recognition and understanding, human often seek to segment an image with semantically meaningful partitions. This is particularly difficult since it requires huge amount of higher level prior information and rules. Currently there has not been any general solution approaching human level segmentations that can be easily achieved by a normal person.

A wide range of early literatures try to produce results where pixels within the same partitioned segments are visually similar. Canonical methods include segmentation by

(a) (b)

Fig. 1. Segmentations of *Toco Toucan*. (a) Result obtained by EGS. (b) Result obtained by our proposed method

maximum a posteriori estimation of Markov random field [1], the family of graph partitioning methods related to spectral graph theory (normalized cut, average cut) [4, 5], arbitrary shaped clustering with mode seeking algorithms [6, 7, 8, 9], region merging [10] and graph-based segmentation methods [11]. But even low level segmentation proves to be hard, as there exist various interference such as large intra-object variance, camouflage, noise, and shading and highlight variance. Segmentation, also, has long been regarded as one of the most computation-consuming computer vision problems [3, 8].

We aim at finding a computationally efficient method that generates visually good image partitionings. Segmentation often serves as an image simplification stage in the recognition framework. Many higher level tasks such as object recognition and scene understanding typically require regional support. The key issue considered here is: how to maximize the simplification level while preserving its region labeling accuracy. Even though a more simplified image probably conveys a clearer scene structures, one also risks more with false labeling in the segmentation process. Oversegmentation, on the contrary, is an effective strategy to reduce false labeling, as is adopted by a number of literatures. For references we recommend readers look into works concerning superpixel or the watershed algorithm [12, 13], where the methods are intrinsically featured with oversegmentations. The disadvantage with such strategy, however, is that useful information such

This work has been supported in part by the Research Grants Council (RGC) of the Hong Kong Special Administrative Region, China. (GRF 610109).

as shape and object topology are basically discarded. Another issue is the complexity of the designed method. We argue, that the algorithm should be fast, simple to implement and intuitively easy to understand.

Our work is closely related to EGS, a very simple yet effective graph-based image partitioning method proposed by Felzenszwalb et al [11]. For the problem's tractability, the authors defined the weakest inter-region link as the inter-region difference. This can be problematic, which is pointed out by themselves in the paper. The accompanied problem is over-smoothing due to certain weak object boundaries or camouflage, which happens frequently in real situations. We address this problem by introducing predicate defined with the maximum likelihood (ML) estimation of inter-region dissimilarity and biased thresholding with intra-region variance. We further make a relaxation to balance the segmentation, by introducing mutual volume for the threshold function. Although relaxing the predicate makes the problem intractable, one shall see such relaxation is essential in producing favorable results. Figure 1 shows the segmentation examples using EGS and our proposed method.

The following parts of this paper will give a detailed discussion on the proposed method. In section 2, we introduce the ML edge predicate and the biased intra-region variance. We then discuss the properties related to the proposed methods in the next session. Based on the discussion, we also introduce a relaxation to the original problem. The experimental results are illustrated in section 4. Finally, conclusions are made in the last section.

2. GRAPH-BASED IMAGE SEGMENTATION

Our proposed method belongs to the family of graph-based image segmentations. Suppose $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ represents an undirected graph where $\mathbf{V} = \{v_i | i = 1, ..., N\}$ denotes the set of nodes corresponding to image pixels and $\mathbf{E} = \{(v_i, v_j) | (i, j) \in \mathbf{S}_{8-Neigh}\}$ denotes the neighboring node pairs. The set $\mathbf{S}_{8-Neigh}$ indicates the 8 connectivity of pixel pair (i, j) in the image space and each $(v_i, v_j) \in \mathbf{E}$ is related with a non-negative edge weight $w(v_i, v_j)$ which is the dissimilarity measure of the two nodes. In this paper $w(v_i, v_j)$ is chosen as the Euclidean distance between RGB values of neighboring nodes, normalized by $\sqrt{3}$. Our goal is to find $\mathbf{P}$, a partition of $\mathbf{V}$ such that each component (or region) $\mathbf{C} \in \mathbf{P}$ corresponds to a connected component in a graph $\mathbf{G}' = (\mathbf{V}, \mathbf{E}')$, where $\mathbf{E}' \subseteq \mathbf{E}$. One could see the definition of $\mathbf{P}$ is identical to the one in EGS.

2.1. Proposed pairwise region comparison predicate

We define a novel pairwise region comparison predicate as the boundary estimator. Instead of choosing the weakest inter-region link as the pairwise region difference, one naturally looks into the ML estimation to robustly reject outliers and weak boundary portions. Suppose $\mathbf{C}_1$ and $\mathbf{C}_2 \subseteq \mathbf{V}$ denote components corresponding to two different regions in the image space, the inter-region difference is defined as:

$$Dif(\mathbf{C}_1, \mathbf{C}_2) = \frac{1}{|\mathbf{B}_{C_1,C_2}|} \sum_{(v_i,v_j)\in\mathbf{B}_{C_1,C_2}} w(v_i, v_j), \quad (1)$$

where $\mathbf{B}_{C_1,C_2} = \{(v_i, v_j) | v_i \in \mathbf{C}_1, v_j \in \mathbf{C}_2, (v_i, v_j) \in \mathbf{E}\}$ is the set of inter-region links between $\mathbf{C}_1$ and $\mathbf{C}_2$ and $|\mathbf{B}_{C_1,C_2}|$ denotes the cardinality of $\mathbf{B}_{C_1,C_2}$. If $\mathbf{C}_1$ and $\mathbf{C}_2$ are non-adjacent, then $|\mathbf{B}| = 0$ and $Dif(\mathbf{C}_1, \mathbf{C}_2)$ is defined to be ∞.

We also define the intra-region difference as the ML estimation of the intra-region links:

$$Int(\mathbf{C}) = \frac{1}{|\mathbf{E}_C|} \sum_{(v_i,v_j)\in\mathbf{E}_C} w(v_i, v_j), \quad (2)$$

where $\mathbf{E}_C = \{(v_i, v_j) | v_i, v_j \in \mathbf{C}, (v_i, v_j) \in \mathbf{E}\}$ corresponds to the set of intra-region links. Intuitively, intra-region difference measures the compactness of a certain region. Our definition here differs from [11] in the sense that such formulation prevents sudden large increase of intra-region difference. In fact, one can also consider defining $Int(\mathbf{C})$ with K-largest $Dif(\mathbf{C}_1, \mathbf{C}_2)$, where $\mathbf{C}_1, \mathbf{C}_2 \subseteq \mathbf{C}$ correspond to regions that are previously merged to form $\mathbf{C}$. Such kind of definition is actually compatible with our framework and is a trade off between [11] and our proposed method. In this paper, however, we only adopt the ML estimation of all links as the intra-region difference.

To estimate whether there exist a boundary between regions, we define the pairwise region comparison predicate similar to EGS:

$$D(\mathbf{C}_1, \mathbf{C}_2) = \begin{cases} \text{true} & \text{if } Dif(\mathbf{C}_1, \mathbf{C}_2) > Mint(\mathbf{C}_1, \mathbf{C}_2) \\ \text{false} & \text{otherwise} \end{cases}, \quad (3)$$

where $Mint(\mathbf{C}_1, \mathbf{C}_2)$ is defined as:

$$\begin{aligned} &Mint(\mathbf{C}_1, \mathbf{C}_2) \\ &= \min(Int(\mathbf{C}_1 + \tau(\mathbf{C}_1)), Int(\mathbf{C}_2 + \tau(\mathbf{C}_2))). \end{aligned} \quad (4)$$

Here we introduce the notion of biased thresholding with intra-region difference. Instead of simply choosing $\tau(\mathbf{C}) \propto \frac{1}{|\mathbf{C}|}$, we define:

$$\begin{aligned} \tau(\mathbf{C}) &\propto \left(\frac{Int(\mathbf{C})}{Int(\mathbf{P})}\right)^{\alpha} \frac{1}{f(|\mathbf{C}|)} \\ &= \left[\frac{(\sum_{k=1}^{K} |\mathbf{E}_{C_k}|) \sum_{(v_i,v_j)\in\mathbf{E}_C} w(v_i, v_j)}{|\mathbf{E}_C| \sum_{k=1}^{K} \sum_{(v_i,v_j)\in\mathbf{E}_{C_k}} w(v_i, v_j)}\right]^{\alpha} \frac{1}{f(|\mathbf{C}|)} \end{aligned} \quad (5)$$

where α is the parameter controlling the strength of bias and $f(|\mathbf{C}|)$ is a monotonically increasing function of $|\mathbf{C}|$. With the above formulation we add additional adaptivity to the thresholding function with respect to intra-region difference, which is normalized by the weighted average of intra-region difference of all currently formed regions. The intuition here is to introduce merging bias towards textured regions. Indeed, we observe that other than oversmoothing caused by weak object boundary, EGS also tends to oversegment textured regions due to the simple formulation of thresholding function. Thresholding plays a crucial rule in determining the segmentation quality since it is a compensation for the region statistics estimated by $Int(\mathbf{C})$. One could also interpret our formulation of $\tau(\mathbf{C})$ as scaling it with intra-region difference. A region with a large $Int(\mathbf{C})$ is likely to be textured regions. For such kind of regions we encourage the merging of textures by increasing $\tau(\mathbf{C})$, which leads to the potential increase of $Mint$

2.2. Segmentation algorithm

Suppose $\mathbf{P}_n = \{\mathbf{C}_1, \mathbf{C}_2, ..., \mathbf{C}_K\}$ is the current segmentation state and $\mathbf{P}_{n+1}$ is the next state derived from $\mathbf{P}_n$. Given the pairwise region comparison predicate, we define the following algorithm to perform segmentation:

1. Initialize $\mathbf{P}_0$ with each component $\mathbf{C}_k$ representing a single image pixel.
2. Repeat step 3 until all pairwise region adjacency links $(\mathbf{C}_{k_1}, \mathbf{C}_{k_2})$ are considered.
3. Construct $\mathbf{P}_{n+1}$ from $\mathbf{P}_n$ as follows. If there exists any link $(\mathbf{C}_{k_1}, \mathbf{C}_{k_2})$ that has not been considered, find the smallest unconsidered $Dif(\mathbf{C}_{k_1}, \mathbf{C}_{k_2})$. If $Dif(\mathbf{C}_{k_1}, \mathbf{C}_{k_2}) \leq Mint(\mathbf{C}_{k_1}, \mathbf{C}_{k_2})$, merge $\mathbf{C}_{k_1}$ and $\mathbf{C}_{k_2}$. Denote the newly formed region as $\mathbf{C}'$. For any region $\mathbf{C}_{k_3}$ adjacent to both $\mathbf{C}_{k_1}$ and $\mathbf{C}_{k_2}$, delete redundant links by merging $(\mathbf{C}_{k_3}, \mathbf{C}_1)$ and $(\mathbf{C}_{k_3}, \mathbf{C}_2)$ into $(\mathbf{C}_{k_3}, \mathbf{C}')$. Update inter-region difference:

$$Dif(\mathbf{C}_{k_3}, \mathbf{C}') = \frac{1}{|\mathbf{B}_{C_{k_3},C'}|} \sum_{(v_i,v_j)\in\mathbf{B}_{C_{k_3},C'}} w(v_i, v_j)$$
$$= \frac{Dif(\mathbf{C}_{k_3}, \mathbf{C}_1)|\mathbf{B}_{C_{k_3},C_{k_1}}| + Dif(\mathbf{C}_{k_3}, \mathbf{C}_2)|\mathbf{B}_{C_{k_3},C_{k_2}}|}{|\mathbf{B}_{C_{k_3},C_{k_1}}| + |\mathbf{B}_{C_{k_3},C_{k_2}}|}$$

Otherwise, $\mathbf{P}_{n+1} = \mathbf{P}_n$. Mark $(\mathbf{C}_{k_1}, \mathbf{C}_{k_2})$ as unconsidered.
4. Return $\mathbf{P}$ and output the segmented image.

2.3. Related properties

We analyze some of the properties related to the above algorithm. To tract the segmentation quality, we need to define the fineness and coarseness of the produced result. Here the definition is identical to EGS:

Definition 1 *A partition $\mathbf{P}$ is too fine if there exist some pair of regions $\mathbf{C}_1, \mathbf{C}_2 \in \mathbf{P}$ for which there is no evidence for a boundary between them.*

Definition 2 *$\mathbf{P}'$ is a refinement of $\mathbf{P}$ if and only if $\forall \mathbf{C}_i \in \mathbf{P}', \exists \mathbf{C}_j \in \mathbf{P}, \mathbf{C}_i \subseteq \mathbf{C}_j$. $\mathbf{P}'$ is a proper refinement of $\mathbf{P}$ when $\mathbf{P}' \neq \mathbf{P}$.*

Definition 3 *A partition $\mathbf{P}$ is too coarse when there exists a proper refinement of $\mathbf{P}$ that is not too fine.*

More details and discussions about the definitions can be found in [11] and we will not extend the discussion here. With the above definitions we are now able to evaluate the image partitionings produced by our proposed algorithm:

Lemma 1 *For any region $\mathbf{C}_{i_m} \in \mathbf{P}_m$, the weakest inter-region difference increases monotonically if $\mathbf{C}_{i_m}$ has not been merged in subsequent operations. In other words, $\min_{j_n}(Dif(\mathbf{C}_{i_n}, \mathbf{C}_{j_n})) > \min_{j_m}(Dif(\mathbf{C}_{i_m}, \mathbf{C}_{j_m}))$ if $n > m$ and $\mathbf{C}_{i_n} = \mathbf{C}_{i_m}$.*

Proof: Lemma 1 is a direct result from the inter-region difference updating rule in algorithm step 3. Suppose two regions $\mathbf{C}_{j_1}$ and $\mathbf{C}_{j_2}$ are both adjacent to $\mathbf{C}_i$ and are merged to form $\mathbf{C}_{j'}$. Since the inter-region difference is defined as the set of all the graph edges going across two distinct regions, inter-region difference can be updated by taking the weighted average of $Dif(\mathbf{C}_i, \mathbf{C}_{j_1})$ and $Dif(\mathbf{C}_i, \mathbf{C}_{j_2})$. One thus have: $Dif(\mathbf{C}_i, \mathbf{C}_{j'}) \geq min(Dif(\mathbf{C}_i, \mathbf{C}_{j_1}), Dif(\mathbf{C}_i, \mathbf{C}_{j_2}))$. This leads to the proof of Lemma 1.

Lemma 2 *For any considered pair of regions where the regions are not merged, at least one of them will be in the final segmentation.*

Proof: Without loss of generality, suppose $\mathbf{C}_i$ and $\mathbf{C}_j$ is the pair of regions one is currently considering and $Dif(\mathbf{C}_i, \mathbf{C}_j)$ is the weakest inter-region difference of $\mathbf{C}_i$. Assume:

$$\begin{aligned} Dif(\mathbf{C}_i, \mathbf{C}_j) &> Mint(\mathbf{C}_i, \mathbf{C}_j) \\ &= \min(Int(\mathbf{C}_i + \tau(\mathbf{C}_i)), Int(\mathbf{C}_j + \tau(\mathbf{C}_j))) \\ &= Int(\mathbf{C}_i + \tau(\mathbf{C}_i)). \end{aligned}$$

Since inter-region difference is considered in a non-decreasing order and according to Lemma 1, we have: $Dif(\mathbf{C}_i, \mathbf{C}_{j'}) \geq Dif(\mathbf{C}_i, \mathbf{C}_j), \forall \mathbf{C}_{j'} \neq \mathbf{C}_j$. In addition:

$$\begin{aligned} Mint(\mathbf{C}_i, \mathbf{C}_{j'}) &= \min(Int(\mathbf{C}_i + \tau(\mathbf{C}_i)), Int(\mathbf{C}_{j'} + \tau(\mathbf{C}_{j'}))) \\ &\geq Int(\mathbf{C}_i + \tau(\mathbf{C}_i)) \\ &= Mint(\mathbf{C}_i, \mathbf{C}_j) \end{aligned}$$

In other words, $Dif(\mathbf{C}_i, \mathbf{C}_{j'}) > Mint(\mathbf{C}_i, \mathbf{C}_{j'})$ and no merging will happen to $\mathbf{C}_i$ in subsequent operations. Thus we have proved Lemma 2

Theorem 1 *The segmentations produced by the proposed algorithm is always not too fine.*

Proof: The produced segmentation is too fine if there exist some pair of regions, say $\mathbf{C}_i$ and $\mathbf{C}_j$, whose comparison predicate does not hold. Without loss of generality, suppose $\mathbf{C}_{j'} \subseteq \mathbf{C}_j$, $Dif(\mathbf{C}_i, \mathbf{C}_{j'}) = Dif(\mathbf{C}_i, \mathbf{C}_j)$ is considered in step 3 and $Int(\mathbf{C}_i) + \tau(\mathbf{C}_i) \leq Int(\mathbf{C}_{j'}) + \tau(\mathbf{C}_{j'})$.

Since $\mathbf{C}_i$ and $\mathbf{C}_j$ are not merged, the corresponding predicate $D(\mathbf{C}_i, \mathbf{C}_{j'})$ considered in step 3 must be true. However, by Lemma 2 we know at least one of the two regions will be a component of the final segmentation. In other words, $Dif(\mathbf{C}_i, \mathbf{C}_j) > Mint(\mathbf{C}_i, \mathbf{C}_j)$ also holds true, which is a contradiction.

Theorem 1 states that the boundaries estimated by our proposed boundary estimator tends to be true object boundaries. In other words, the estimated boundaries are reliable. On the other hand, we are not able to guarantee that the method could find all true object boundaries since segmentations produced by the algorithm is possible to be too coarse.

Does the algorithm produces results with more serious oversmoothing than EGS? Actually not. In fact, our proposed predicate makes a stronger boundary estimator than EGS in the sense that one has the opportunity to correct weak object boundary by averaging it with strong parts of the object boundary. We can interpret the comparison of our method and EGS in the following way. In EGS, the merging predicate always holds since the inter-region difference that causes the merging remains the same no matter how the merged regions grow by merging other regions. What makes a difference is our definition of inter-region difference: the inter-region difference that causes the merging always grows larger as more regions are merged to the previously merged regions. In other words, our definition of inter-region difference makes one "realize" there is a chance that the merged regions should actually be separated. The same error could probably also happen to EGS but one simply does not "realize" the previously made error.

3. RELAXATION WITH MUTUAL VOLUME

We implement the segmentation algorithm described in section 2.2 and compare its segmentation results with segmentations obtained by EGS. Figure2 illustrates the original images and the comparison of results obtained by different algorithms. The images are first smoothed by a gaussian kernel with $\sigma = 0.8$ and both algorithms are performed on the grid graph whose definition is described in the beginning of Section 2. For EGS we set $\tau(\mathbf{C}) = \frac{300}{|\mathbf{C}|}$, where the parameter $k = 300$ is the recommended value that has been adopted in experiments of [11]. Since our proposed predicate makes a stricter boundary estimator, we relax the constraint by setting a larger k value and suppress the increase of $f(|\mathbf{C}|)$ as $|\mathbf{C}|$ increases. Specifically, we set $k = 400$. $f(|\mathbf{C}|)$ is defined as:

$$f(|\mathbf{C}|) = \begin{cases} \sqrt{|\mathbf{C}|} & \text{if } |\mathbf{C}| < 4000 \\ \sqrt[3]{|\mathbf{C}|} - \sqrt[3]{4000} + \sqrt{4000} & \text{otherwise} \end{cases}.$$

In addition, the bias strength controlling parameter α is set to 1.

From figure2 we can see our proposed method produces more favorable results by adopting a boundary estimator stricter than EGS. Moreover, adding more adaptivity to the thresholding function which compensates region statistics estimation further helps to improve the segmentation quality. We also observe, however, that the algorithm still tend to over-segment where there exist strong textures and noises. One can infer that the fundamental reason is the formulation of a region's thresholding function being independent with respect to other region volumes. In fact, such formulation is indispensable for the problem's tractability, but is not always necessary in producing good segmentations. Our strategy for penalizing oversegmentation is to introduce pairwise mutual region volume - a value defined as the geometric mean of two adjacent regions' volume - to substitute each single region volume in determining the thresholding value. Even though the relaxation leads to intractability of the segmentation problem, one shall see such approach is actually very powerful in producing good segmentations. To implement the above relaxation, we do not need to change the algorithm except simply making a small modification to the minimum intra-region difference:

$$\begin{aligned} &Mint(\mathbf{C}_1, \mathbf{C}_2) \\ &= \min(Int(\mathbf{C}_1 + \tau(\mathbf{C}_1, \mathbf{C}_2)), Int(\mathbf{C}_2 + \tau(\mathbf{C}_1, \mathbf{C}_2))), \end{aligned} \tag{6}$$

where

$$\begin{aligned} &\tau(\mathbf{C}_1, \mathbf{C}_2) \\ &= \left[\frac{(\sum_{k=1}^{K} |\mathbf{E}_{C_k}|) \sum_{(v_i,v_j)\in \mathbf{E}_C} w(v_i, v_j)}{|\mathbf{E}_C| \sum_{k=1}^{K} \sum_{(v_i,v_j)\in \mathbf{E}_{C_k}} w(v_i, v_j)}\right]^{\alpha} \frac{1}{f(\sqrt{|\mathbf{C_1}||\mathbf{C_2}|})} \end{aligned} \tag{7}$$

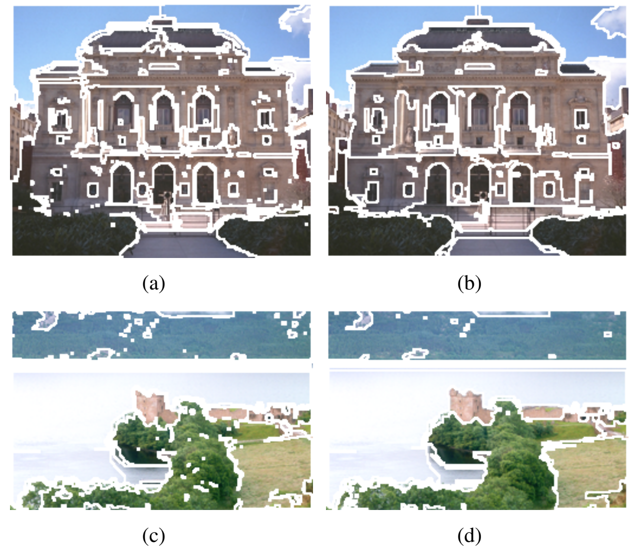

Fig. 2. Segmentations of *Opera* and *Loch Ness*. (a) Result obtained by EGS. (b) Result obtained by the proposed algorithm

Table 1. Quantitative evaluation

Image	F'(I)			Q(I)		
	EGS	Ours	QS	EGS	Ours	QS
Brandy Rose	2.71	**0.44**	1.82	5.24	**2.56**	3.12
Butterfly	0.72	**0.08**	0.23	1.13	**0.51**	0.76
Cow	4.37	**0.23**	0.6	6.14	**1.23**	1.75
Flowers	8.33	**0.73**	3.13	8.7	**3.6**	4.53
Frangipani1	2.08	**0.3**	0.62	3.67	2.03	**1.89**
Frangipani2	4.85	**0.64**	1.36	6.4	3.4	**3.1**
Kids	2.68	**0.29**	2.51	5.49	**2.35**	2.86
Lake	0.48	**0.05**	0.05	1.5	0.63	**0.52**
Loch Ness	2.3	**0.12**	0.34	3.18	**0.81**	1.14
Mountain	1.95	**0.18**	0.18	3.25	1.67	**0.64**
Opera	2.44	0.23	**0.16**	3.35	1.19	**1.09**
Peppers	4.48	**0.66**	0.78	8.02	3.55	**2.24**
Red Building	2.19	**0.66**	0.1	2.47	2.27	**0.86**
Skyline Arch	5.13	**0.78**	2.9	6.01	**2.95**	4.2
Toco Toucan	1.88	**0.24**	0.67	3.71	**1.66**	1.8
Water Lilies	10.4	**0.77**	4.59	10.2	**3.68**	4.55
Hand	6.45	**0.36**	1.2	7.58	**1.62**	2.03
Horse	5.55	**0.29**	1.01	7.03	**2.1**	2.56
Average	3.83	**0.39**	1.24	5.17	**2.1**	2.2

4. EXPERIMENTAL RESULTS

We implement the algorithm with the relaxation proposed in section 3 to segment a number of test color images and evaluate its performance. The corresponding parameters are identical to those in section 3. EGS with parameter k equal to 300 was implemented for the purpose of comparison. We also compare our results with quick shift (QS), a fast and effective mode seeking algorithm recently proposed for applications of image segmentation. We run the quick shift algorithm with the VLFeat Matlab package which is kindly available at http://www.vlfeat.org/. The parameters $ratio$, $kernelsize$ and $maxdist$ are respectively set to 0.5, 10 and 30. Results produced by the above three methods are partially illustrated in figure 3. One could observe that our proposed method tend to generate most favorable segmentations with less over-segmentions and more continuously preserved major object boundaries, which can potentially lead to less false labeling and beneficial gains in subsequent operations.

We also perform quantitative evaluation to the produced results. Details about the benchmarks can be found in [2] and the benchmarks have been well designed for evaluating unsupervised segmentations. The evaluation of segmentation results are illustrated in table 1. Results indicate our method outperforms the other two methods on the adopted two benchmarks.

5. CONCLUSIONS

In this paper, we propose an effective segmentation method that generates favorable partitioned results, as is supported by natural image segmentation tests. Our approach is related to graph-based partitioning method, which is both computationally efficient and intuitively simple. In fact, the method can be easily extended to the framework with superpixels where the features for grouping can be designed in a much more versatile and sophisticated way, leading to better segmentation results or favorable biases in specific tasks. Further improvements and discussions will be included in our future works.

6. REFERENCES

[1] S. Geman and D. Geman. "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images." *IEEE Trans. PAMI*, 1984.

[2] H. Zhang, J. E. Fritts and S. A. Goldman. A co-evaluation framework for improving segmentation. In *Proceedings of the SPIE*, vol. 5809, pp.420-430, 2005.

[3] Y. Boykov, O. Veksler and R. Zabih. "Fast Approximate Energy Minimization via Graph Cuts." *IEEE Trans. PAMI*, 2001.

[4] J. Shi and J. Malik. "Normalized cuts and image segmentation." *IEEE PAMI*, 2000.

[5] F.R.K. Chung. "Spectral graph theory." Am. Math. Soc., 1997.

[6] D. Comaniciu and P. Meer. "Mean shift: A robust approach toward feature space analysis." *IEEE Trans. PAMI*, 2002.

[7] Y. A. Sheikh, E. A. Khan and T. Kanade. "Mode-seeking by Medoidshifts." In *ICCV*, 2007.

[8] A. Vedaldi and S. Soatto. "Quick shift and kernel methods for mode seeking." In *ECCV*, 2008.

[9] A. Vedaldi and S. Soatto. "Really quick shift: Image segmentation on a GPU." In *Workshop on Computer Vision using GPUs, held with the ECCV*, 2010.

[10] R. Nock and F. Nielsen. "Statistical region merging." *IEEE Trans. PAMI*, 2004.

[11] P. Felzenszwalb. "Efficient graph-based image segmentation." *IJCV*, 2004.

[12] S. Beucher and F. Meyer. "The morphological approach to segmentation: The watershed transformation." In *Mathematical Morphology in Image Processing*, 1993.

[13] X. Ren and J. Malik. "Learning a classification model for segmentation." In *ICCV*, 2003.

Fig. 3. Segmentation Results. The first column contains the original test images. The second and third column correspond to results obtained by EGS. The fourth and fifth column are segmentations produced our proposed method. The last column are segmentations produced by quick shift. The corresponding test images from the first row to the last row are respectively *Cow*, *Hand*, *Toco Toucan*, *Lake*, *Loch Ness*, *Peppers*, *Opera* and *Red Building*.