

Research Article

KmsGC: An Unsupervised Color Image Segmentation Algorithm Based on K -Means Clustering and Graph Cut

Binmei Liang^{1,2} and Jianzhou Zhang¹

¹ College of Computer Science, Sichuan University, Chengdu, Sichuan 610065, China

² College of Mathematics and Information Science, Guangxi University, Nanning, Guangxi 530004, China

Correspondence should be addressed to Binmei Liang; liangbinmei@sina.com

Received 26 August 2013; Revised 17 March 2014; Accepted 2 April 2014; Published 12 May 2014

Academic Editor: Gerhard-Wilhelm Weber

Copyright © 2014 B. Liang and J. Zhang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

For unsupervised color image segmentation, we propose a two-stage algorithm, KmsGC, that combines K -means clustering with graph cut. In the first stage, K -means clustering algorithm is applied to make an initial clustering, and the optimal number of clusters is automatically determined by a compactness criterion that is established to find clustering with maximum intercluster distance and minimum intracluster variance. In the second stage, a multiple terminal vertices weighted graph is constructed based on an energy function, and the image is segmented according to a minimum cost multiway cut. A large number of performance evaluations are carried out, and the experimental results indicate the proposed approach is effective compared to other existing image segmentation algorithms on the Berkeley image database.

1. Introduction

Unsupervised color image segmentation plays a key role in various image processing and computer vision applications, such as medical image analysis [1], image retrieval [2], image editing [3], and pattern recognition [4]. Although the reported performance of image segmentation algorithms has been improved significantly these years, it is by general consent that there is still a long way to reach a high-quality algorithm for unsupervised image segmentation. Among the approaches for color image segmentation, edge-based approaches [5] detect region boundaries based on the assumption that the pixel properties should change abruptly between different regions. A common issue of these approaches is the generation of disconnected edges which leads to undersegmentation. Region-based approaches [6] employ a set of pixels as starting seeds to grow into homogenous grouping of pixels. However, the efficiency of these segmentation approaches relies heavily on the choice of seeds. Watershed-based approaches [7, 8] use gradient information and morphological markers for the segmentation. While these approaches achieved meaningful

segments, they are computationally intensive. Clustering-based approaches [9, 10] capture the global characteristics of the images by calculating the image features, usually color or texture, to efficiently segregate data. In [9], the authors proposed a feature space analysis and clustering-based mean-shift segmentation algorithm called MS algorithm. It uses kernel density estimation to model feature space data and reaches the local minimum or maximum by finding the zero of the gradient of the density function so as to obtain the clustering centers. A notable advantage of this algorithm is its modularity; however, this algorithm needs a complicated feature selection process, and it is not practically easy to choose appropriate parameters for various images. And the segmentation result is somewhat sensitive to resolution parameters h_s and h_r . In [10], a CTM algorithm is presented, and it segments the images by clustering the texture features. It models the distribution of the texture features with a mixture of Gaussian distributions. The clustering algorithm derives from a lossy data compression. CTM algorithm compresses feature vectors or windows extracted around all pixels by partitioning them into clusters. As a result, the final coding length is highly redundant due to severe

overlap between windows of adjacent pixels, which could lead to incorrect segmentation. CTM algorithm encodes the membership of the pixel by a Huffman coding, but it does not take account of spatial adjacency of pixels nor smoothness of boundaries, which cannot ensure correct and smooth boundary. In practice, it is hard to choose proper parameters, and the segmentation results are sensitive to the parameters. Graph-based approaches [11, 12] construct a weighted graph with limited vertices and edges, and the graph is partitioned into multiple components according to the minimization of a cost function to achieve the segmentation. So far, many graph-based approaches have been developed for color image segmentation, and their drawbacks generally lie in the high computation complexity. Hybrid approaches [13–15] have become a noticeable developing direction over the years. They usually mix edge and region information together with other image features or combine two segmenting techniques to achieve optimal results. The approach JSEG in [13] quantizes the color to several classes to form a class-map and then use region growing technique to obtain the segments, but the parameters are hard to be specified. The work [15] is a hybrid approach which employs the vector-based color gradient method and Otsu's automatic threshold to perform a dynamic threshold-based segmentation. All these approaches address some of the many drawbacks in color image segmentation, such as complex computing models, computation expensiveness, and sophisticated parameters.

Apart from these approaches, graph cut based techniques emerged as increasingly useful tools for energy minimization problems producing provably high-quality solutions in practice, and they have been successfully applied to a wide range of problems in computer vision in recent years. In image segmentation context, since Boykov and Jolly first demonstrated how to use binary graph cuts to build efficient object extraction tools for image segmentation in [16], many graph cut based extensions have been proposed, for example, geo-cuts, grab-cuts, lazy snapping, obj-cuts, and so forth. Generally speaking, these works belong to binary graph cuts formulation or interactive solution which means that the user should place seeds interactively. In [17], the authors present a graph cut based unsupervised image segmentation solution. This algorithm incorporates data cost, smoothness cost, and label cost in an energy function and calculates the optimum energy by binary graph cuts formulation. It uses the minimum description length principle to obtain an unsupervised segmentation. Current graph cut based image segmentation techniques are mainly used in binary image segmentation [16, 18] and interactive image segmentation [19–22].

In this work, we aim at making use of the advantages of clustering technique and graph cut optimization technique, combining them together, and propose a new unsupervised color image segmentation solution. We solve the color image segmentation problem via multiway graph cuts and establish a compactness criterion to achieve an unsupervised, fast, and effective segmenting solution.

The rest of the paper is organized as follows. In Section 2, we review the graph cut framework and the K -means clustering technique and describe how to determine the number

of clusters. The new segmentation approach is presented in Section 3. Experimental results, discussion, and conclusion are presented in Sections 4 and 5, respectively.

2. Background

2.1. Image Segmentation Based on Graph Cut. An image can be defined by a pair (P, I) consisting of a finite discrete set $P \subset Z^d$ ($d > 0$) of points (pixels in Z^2 , voxels in Z^3 , etc.) and a function I that maps each point $p \in P$ to a value $I(p)$ in some value space. Usually, P corresponds to a square. For an image, we can construct a nonnegative edge weighted graph $G = (v, \varepsilon, c)$ consisting of a set of vertices $v = P \cup L$, a set of edges ε , and a positive weighting function $c : v^2 \rightarrow R^+$ defining the edge capacity.

We distinguish between two special vertices of v : the multiple terminal vertices $k \in L = \{1, 2, \dots, k\}$, where L is a label set, and pixel vertices $p \in P$. Every pixel vertex has an n -link to its four neighbors and also a t -link to every terminal vertex. The set of edges ε consists of all n -links and t -links.

In [23, 24], Boykov et al. showed that the image segmentation problem can be naturally formulated in terms of energy minimization, and we use Potts energy in the form of

$$E(L) = \sum_{p \in P} D_p(L_p) + \sum_{(p,q) \in N} V_{p,q}(L_p, L_q), \quad (1)$$

where N is the set of interacting pairs of pixels, $D_p(\cdot)$ is a data penalty term, $V_{p,q}$ is a discontinuity penalty term, and $L = \{L_p \mid p \in P\}$ is a labeling of image P .

For the energy minimization problem, there are some methods for the optimality guarantees in certain cases. For example, dynamic programming can be used for a few one-dimensional energy function minimization, such as snakes. Mean field annealing can deduce the minimum of the energy by estimating the partition function. But computing of the partition function is computationally intractable. Graduated nonconvexity is a kind of continuation method, but the quality of the output may not be known except for certain cases (see [25] for details). For the local minimum of energy minimization, lots of methods can be used for the continuous energy as described in [26], and Euler equations are often used to guarantee a local minimum. Another alternative is to use discrete relaxation labeling methods in which combinatorial optimization is converted to continuous optimization with linear constraints and then gradient descent is used to give the solution. Graph cut is a powerful technique for solving many computer vision problems and we apply it to solve the energy minimization in this work. According to [23, 24], the global minimum of energy (1) can be computed by computing a minimum cost multiway cut on an appropriately constructed graph as Figure 1 [23].

2.2. K -Means Clustering Technique. Clustering is a process of classifying data items into similar groupings. K -means is undoubtedly the most widely used clustering algorithm for its simplicity, efficiency, and empirical success. It generates clusters by minimizing the sum of distances between each data point and its cluster center. There are many ways for

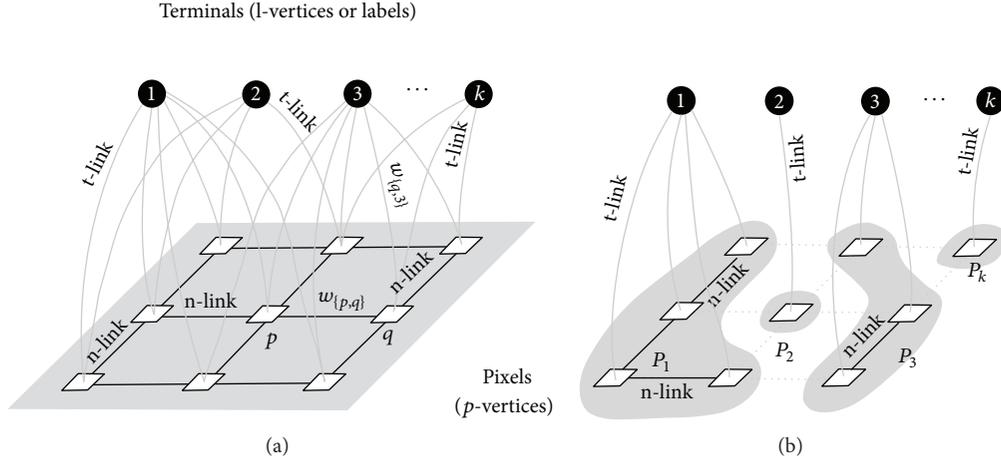


FIGURE 1: (a) A graph $G = (v, \epsilon)$ with multiple terminal vertices $L = \{1, 2, \dots, k\}$. (b) A multiway cut on G corresponding to a unique partition (labeling) of image pixels.

the distance measuring (e.g., squared Euclidean distance, hamming distance, and cosine dissimilarity), and we use the most popular Euclidean distance in this work which forms the SSE (sum-squared-error) criterion:

$$\text{SSE} = \sum_{i=1}^k \sum_{x_j \in P_i} \|x_j - c_i\|_2^2, \quad (2)$$

where $\|\cdot\|_2$ denotes the Euclidean norm, C_i is the centroid of cluster P_i , and k is the number of clusters.

K-means algorithm starts iteration with initial cluster centers and in the iteration each data item is assigned to the nearest cluster center based on the Euclidean distance and after updating the new cluster centers, the data item is reassigned until no change of the new cluster centers.

2.3. How to Determine the Number of Clusters. Determining the optimal number of clusters is one of the most difficult problems for clustering-based image segmentation methods. Most methods cast this problem into the model selection [27]. Minimum description length (MDL) principle is a well-known regularization solution for this problem, first introduced by Rissanen in 1978. MDL principle is based on a simple idea that the best way to capture regular features in data is to construct a model in a certain class which permits the shortest description of the data and the model itself. After formulating a minimum description length criterion of the data, the basic MDL based algorithm starts with a large number of clusters and gradually merges the clusters until the MDL criterion does not decrease. However, the validity of a MDL selection criterion depends on properties of the underlying coding scheme or, more precisely, the resulting description lengths [28]. Besides, formulating an accurate measure of the coding length is hard, and it is difficult to decide the optimal stop time of the merging process. In our approach, the clustering algorithm is run with different cluster number, and the optimal cluster number is chosen based on a compactness criterion. This means calculating

the values of the compactness criterion in the process of clustering to reach a simple and efficient solution.

3. The Proposed Approach

In this part, we proposed a new unsupervised color image segmentation solution. As the number of clusters is necessary for the K-means clustering, we should first determine the number of clusters.

3.1. Determine the Number of Clusters. We know that “minimum intracluster variance and maximum intercluster distance” lead to compact clusters [29] and based on this principle we determine the optimal cluster number. Below we establish a compactness criterion based on this principle to find compact clustering so as to determine the optimal number of clusters.

First, we want to find compact clustering with “minimum intracluster variance.” An intracluster distance measure is defined as the average distance between the data items and their cluster centers within clusters, and we want it to be as small as possible. It can be defined as

$$\text{intra} = \frac{1}{N} \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|^2, \quad (3)$$

where $\|\cdot\|$ denotes the Euclidean norm, k is the number of clusters, N is the number of pixels in the image, and μ_i is the cluster center of cluster C_i . We obviously want to minimize this measure.

Next, we use an intercluster distance measure, the distance between clusters, to describe compact clustering with “maximum intercluster distance,” and we want it to be as large as possible. We calculate it as the distance between cluster centers and take the minimum of this value, defined as

$$\text{inter} = \min \left(\|\mu_i - \mu_j\|^2 \right), \quad (4)$$

where $i = \{1, 2, \dots, k-1\}$ and $j = \{i+1, \dots, k\}$. Obviously, we want to maximize this measure.

Since the minimum value of intrameasure and the maximum value of intermeasure lead to compact clustering, we combine them and define a compactness criterion as

$$CP = \frac{\text{intra}}{\text{inter}}. \quad (5)$$

To obtain compact clusters, we obviously want to minimize the CP criterion, and the minimum value of this criterion corresponds to the best clustering.

3.2. The Proposed Approach. Based on the work above, we present a new unsupervised color image segmentation approach (referred to as KmsGC) which consists of two stages. In the first stage, we give the image pixels initial partitioning clustering by K -means algorithm, and we use the technique described above to automatically determine the optimal number of clusters. In the second stage, we optimize the initial pixel partition by an energy formulation. We construct a weighted graph in terms of the energy function and then compute a minimum cost multiway cut using the max flow/min cut algorithm.

In the first stage, we do clustering with K -means algorithm and determine the optimal number of clusters using the compactness criterion in (5). There are many published methods for determining the optimal number of clusters, such as Akaike information criterion (AIC), minimum description length (MDL) [26], minimum message length (MML), Bayes information criterion (BIC), and gap statistics [30–32]. Here, we apply a simple and effective approach described in [30], which experimentally works well with the proposed framework. The process includes executing an iteration from $k = 2$ to K_{\max} , where K_{\max} represents an upper limit of the number of clusters. In the iteration, the cluster with larger inner variance is partitioned into two subclusters using K -means algorithm and then a new compactness criterion value of this k is calculated. After the iteration, we take the cluster number with the smallest value of compactness criterion as the optimal k .

The inner cluster variance is defined as

$$\sigma_{ij}^2 = \frac{1}{N_i} \sum_{x \in C_i} (x - \mu_{ij})^2, \quad (6)$$

where $i = \{1, 2, \dots, k\}$ and $j = \{1, 2, 3\}$, σ_{ij} is the j th components for cluster C_i in the color space, N_i is the number of pixels in cluster C_i , and x is the vector representing each pixel's red, green, and blue components, respectively. μ_{ij} is the cluster center of C_i of the j th components. We get σ_i , the variance of cluster C_i , by taking the average of the three components of σ_{ij} . The presented algorithm is implemented in RGB color space.

In the second stage, we address the construction of energy function in our application as follows. We firstly deal with the data penalty term, the first term in (1). In our use, the data penalty term $D_{pi}(p)$ indicates the cost for allocating the pixel p to the cluster i . From the above analysis, it is obvious that our goal is to obtain compact clusters, which means we want

every pixel to belong to the closest cluster center. Therefore, we set data penalty term as the distance between the pixel p and the cluster center:

$$D_{pi}(p) = (\|x_p - \mu_i\|)^2, \quad (7)$$

where $\|\cdot\|$ is the Euclidean norm, x_p ($p = 1, 2, \dots, N$) is the characteristics of pixel p , here we use its RGB value, and μ_i ($i = 1, 2, \dots, k$) is the cluster center of cluster C_i , where k is the number of clusters.

Secondly, we describe the discontinuity penalty term, the second term in (1). As [23] points out, a common constraint for label-assigning is that the label should vary smoothly almost everywhere while preserving sharp discontinuities that may exist at object boundaries, so discontinuity penalty term should penalize neighboring pixels having different labels. We want it providing smooth boundaries for the segmentation, so we set it as

$$V_{pq}(l_1, l_2) = V(l_1, l_2) * W_{pq}, \quad (8)$$

where $V(\cdot)$ is the spatially unvarying smoothness cost and W_{pq} is the spatially varying cost. $V(l_1, l_2)$ indicates the cost of assigning neighboring pixels with l_1 and l_2 ($l_1, l_2 \in L = \{1, 2, \dots, k\}$ where L is a label set) and this cost is spatially invariant. Therefore, the spatially unvarying smoothness cost is set as a k by k matrix where each element value equals

$$V(l_1, l_2) = \begin{cases} 1 & \text{if } l_1 \neq l_2 \\ 0 & \text{otherwise.} \end{cases} \quad (9)$$

We set the spatially varying smoothness cost as the filtering of the image according to each RGB channel and take the maximum of the three values as the final value.

For the energy function minimization algorithm, as the global energy minimization needs enormous computational costs, local minimum becomes a desirable option. Standard moves technique is often used for calculating the local minimization, which changes a pixel's label in a time. Many methods use standard moves, such as iterated conditional modes (ICM) [33] and simulated annealing [34]. ICM chooses the largest decrease of the energy function until it converges to a local minimum. Simulated annealing is an outstanding technique which can optimize arbitrary energy functions, but it needs an exponential time for this. A weak point of standard moves is that it cannot decrease the energy by changing a single pixel's label at a local minimum, which brings low quality solution (see [23] for details). In our work, we use expansion moves which can make larger moves to change labels of large sets of pixels simultaneously to get a fast and accurate solution (see [23] for its detailed performance comparing with simulated annealing). Then, the minimum cost multiway cut is computed by the expansion algorithm presented in [23, 35] and finally the segmentation result is obtained based on the cut.

In summary, the proposed KmsGC algorithm can be described as in Algorithm 1.

1st stage. Clustering**Input:** image I **Output:** initial pixel partition.

- (1) initialize K_{\max}
- (2) for $i = 2: K_{\max}$
- (3) k -means($w, 2$) \ \ w is the cluster to be partitioned into 2
- (4) calculate intra, inter, $CP(i)$
- (5) calculate σ_i
- (6) $w = \text{find}(\max(\sigma_i))$
- (7) endfor
- (8) $k = \text{find}(\min(CP(i)))$
- (9) Return the pixel partition corresponding to k .

2nd stage. Image segmentation**Input:** initial pixel partition**Output:** segmented image

- (1) Initialization.
- (2) Determine the data term as (7)
- (3) Determine the smooth term as (8)
- (4) Construct a graph correspond to the energy function
- (5) Compute the minimum cost multi-way cut using the expansion algorithm.
- (6) Return segmented image

ALGORITHM 1: The basic structure of the presented algorithm KmsGC.

4. Experimental Results

In order to evaluate the performance of the proposed algorithm, we design a number of experiments. First, the accuracy of determining the cluster number must be tested and compared. Second, the qualitative performance is to be accessed by comparing the segmented results with human segmentations and other approaches. Third, the segmented results are to be compared against other approaches by means of effective indices to judge the quantitative performance of the proposal.

Experiments are performed on a typical 2 GHz Intel PC with 2 GB RAM, the MATLAB implementation of our algorithm on spatial point sets, standard images, and real color images.

In terms of the computation time, it depends on the size of image, the value of K_{\max} , and the number of clusters. It mainly consists of two parts: the time for clustering including determining the optimal cluster number and the time for graph cut based segmentation. As for a $481 * 321$ real color image and as K_{\max} equals 25, it takes about 2 seconds to do clustering (including automatically determining the optimal number of clusters) and a few seconds to segment the image via graph cut. Below are the three experimental evaluations for the proposed algorithm.

4.1. Evaluation of the Determination of Clusters. We first test the accuracy of the proposal on the cluster number determination. The proposal was tested on various spatial point sets for their ideal cluster numbers are known beforehand and also on the standard images for these images are simple so their segmentation results are straightforward. In the experiments, we found that the minimum value of the compactness criterion occurred at the ideal cluster number

for each spatial point set, which verifies that the optimal cluster number can be determined according to the minimum value of the compactness criterion. Here, we give the results of a point set and a standard image to show how the algorithm works. Figure 2(a) is a point set with obvious six clusters and we apply it to show the determination of cluster number. We set parameter $K_{\max} = 25$ and after running the algorithm we get the result of Figure 2(b). As can be seen in Figure 2, the proposed algorithm obtains the minimum value of compactness criterion when the cluster number is 6, which verifies that the algorithm is indeed to calculate the correct number of clusters in an unsupervised way.

More than ten standard color images are also used to test the performance of cluster number determination. In these tests, we set $K_{\max} = 25$, and the cluster numbers are thought to be correctly determined if the corresponding segmentation results are good from a visual point of view. Here, we provide the segmentation result of a house picture. Figure 3(c) shows that the compactness criterion reaches the minimum value when the cluster number is 4 and the corresponding segmented image is showed in Figure 3(b). From the result, we can see that the algorithm provides nearly perfect segmentation, which suggests that the cluster number is correctly determined.

Below we compare our unsupervised solution with the popular regularization criterion MDL. There are many MDL criterion based segmentation algorithms, such as [10, 17, 36]. For the coherence consideration, we will take [10] as a main comparison.

For the proposed KmsGC, we use a compactness criterion and the energy function to reach compact clustering, which is consistent with the base point of human segmentation. The MDL based algorithm has been used for many unsupervised segmentation solutions and it is objective, but it is

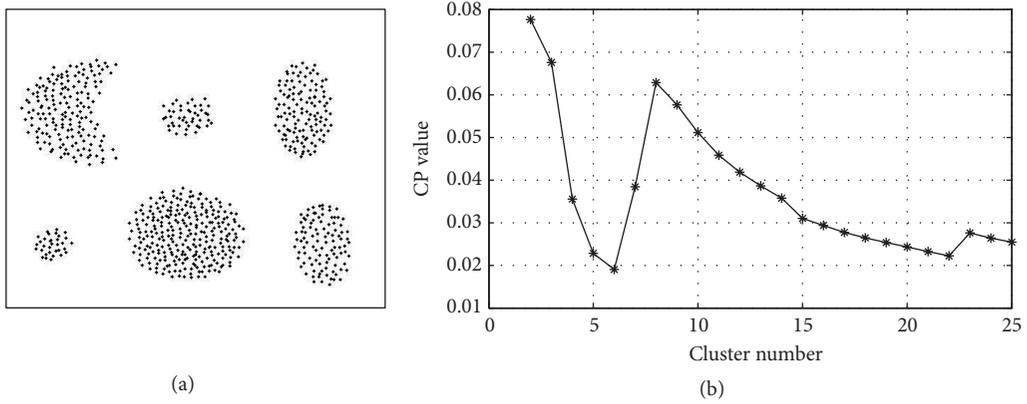


FIGURE 2: (a) A point set and (b) its optimal cluster number determination. In (b), x -axis represents the cluster number whereas y -axis shows the values of the compactness criterion, and $K_{\max} = 25$. The compactness criterion reaches the minimum value at 6 clusters.

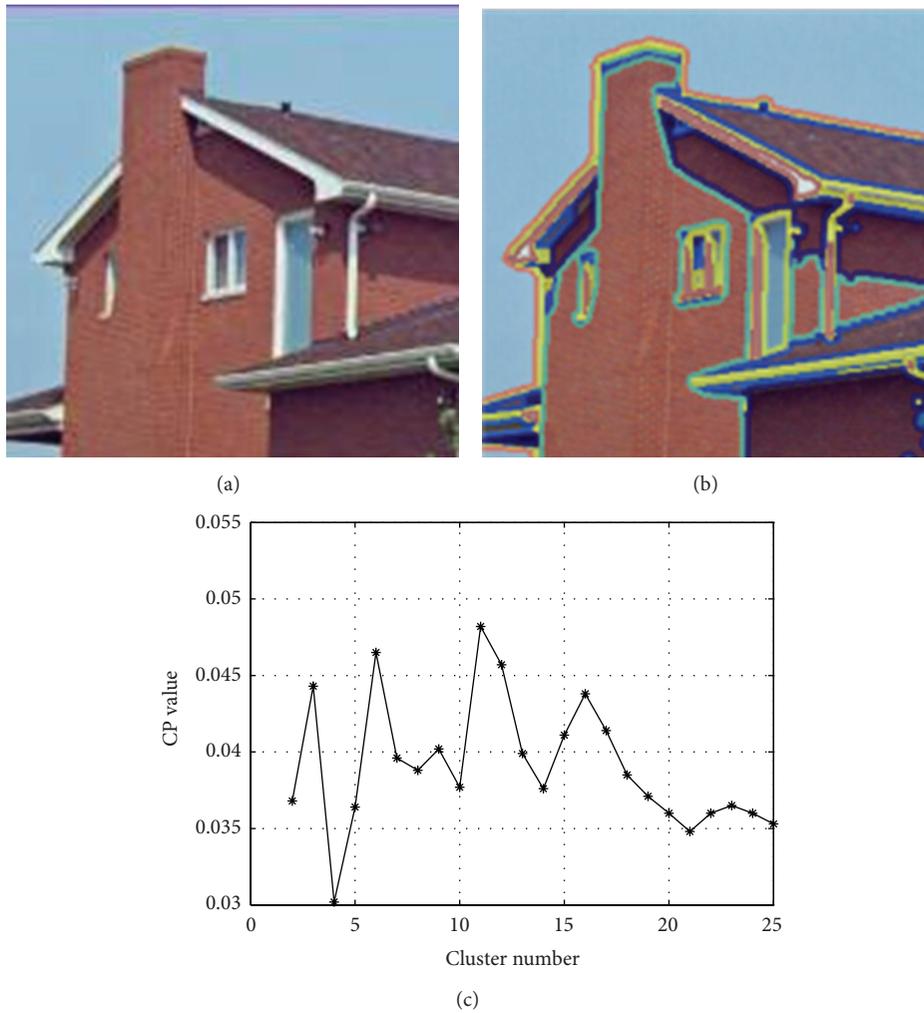


FIGURE 3: (a) An original image, (b) segmented image, and (c) the cluster number determination. In (c), x -axis represents the cluster number, whereas y -axis shows the values of the compactness criterion, and $K_{\max} = 25$. The compactness criterion reaches the minimum value at 4 clusters. (b) is the segmented image with this optimal cluster number.

experimentally difficult to determine the optimal stopping time of the merging, so it may generate oversegmentation or undersegmentation. For example, at the right column of the second row of Figure 13 in [10], the right picture is undersegmented and the middle picture is oversegmented. More examples can be seen in [10]. This defect also can be seen in Figure 14 of [36].

MDL based algorithms usually merge two adjacent segments if this merging decreases the coding length the most until no decrease happens, like [10, 36]. We find that sometimes the MDL based algorithm fails to accurately control which two adjacent segments to be merged in the merging process. For example, in the right picture of the second row of Figure 9 in [10], there are many oversegments of the sky while the land is undersegmented and the grass detail is not segmented. Comparing the 3rd picture of Figure 4 of the proposal, the grass detail is segmented and the oversegments of the sky are less.

Regarding detail preservation, MDL based segmentation sometimes cannot preserve details. For example, in the left picture of the second row in Figure 11 in [10], the man and the car are lost in the final segmentation. In the right picture at the second row of the left column in Figure 10 of [10], the details of some shades and the seacoast are missing. Compare the proposed segmentation in Figure 7. Also observe the eyes and mouths of the woman image in Figure 14 of [36] and Figure 15 of [17].

For the MDL algorithm, if the boundary constraints are not incorporated into the MDL criterion, it may be hard to reach correct and smooth boundaries. Compare our boundary with the results in Figure 14 of [36]. But if it is incorporated, it may be harder to control the segmentation precisely as the coding length gets more complicated. Observe the faces, river, and skin of the second picture in Figure 15 of [17].

4.2. Visual Evaluation. In this section, we visually evaluate the qualitative performance of the segmentation results on the Berkeley image database. This database provides several human manually segmented results for each image and here we employ two of them as the perceptual evaluation references per image. We set the evaluation criterion as: the more similar to these human segmentations, the better the segmented results. We test the proposal on all the 300 images (image size $481 * 321$) of Berkeley image database and some results and the corresponding human segmentations are provided in Figure 4, and some corresponding cluster number determination results are showed in Figure 5. And Figure 8 shows more segmentation results from different categories of the Berkeley images, namely, landscape, objects, water, sky, animals, and people.

From Figure 4, we observe that different regions in the images are well segmented where the contours of the segmented regions match those in the human segmentations, the results are of smooth-boundary and large-region, and the important details in the images have been detected. For example, consider the following.

- (1) In the bear image, the contour of the segmented bear is highly consistent with the first human segment; the segmented boundary of the bear and the bands is smooth. Important details are detected, such as the ears and nose of the bear.
- (2) In the 2nd image, the church is accurately segmented and the boundary is smooth.
- (3) In the third image, the segmented region contours of the stone and the tree are close to those of the human's, and the boundary is smooth. The segmented regions are of large size while the details are preserved.
- (4) In the flower image, note that in the second human segmentation the leaves are segmented while they are not in the first human's. In our result, the petals and the leaves all have been segmented, and the contour of the flower core is correct. The contours of the petals are generally correct, and the leaves are in large segments.
- (5) In the woman image, although the illumination variation exists, the woman's face, neck, and hair are accurately segmented and also the forehead is well segmented, which indicates that the number of clusters is accurate.
- (6) In the 6th image, the objects like the man, fishes, and the sea bottom are well segmented where the number and size of the segmented regions are almost equal to those of human segments, which make them easy to be automatically recognized.
- (7) Observing the result of the 7th image, the distant objects like the tree and hills are accurately segmented where the contour of the tree highly matches those of the human segments, and the boundaries are smooth, which can be a substitute for the human segmentation. In the near area, the cow and the haystack are well segmented in general.

Figure 5 provides the results of the calculation of the compactness criterion on some images of Figure 4, and from the results we obtain the optimal cluster number for those images.

We set the upper limit cluster number $K_{\max} = 25$ for all natural images segmentation, which is approximately more than 2 times of the ideal cluster number of the images and we think it is large enough from a visual point of view. In the experiments, we found that the algorithm works well on large number of the tested images with this fixed K_{\max} , and from our experience it has little change of the results while K_{\max} takes different values within this range.

As the segmented results illustrate that the segmentations are very close to human segmentations, it may suggest that the proposed algorithm on cluster number determination is effective.

We now visually compare the proposal with two clustering-based unsupervised segmentation algorithms: MS [9] and CTM [10].

First compare with the MS algorithm. Observing the experimental results of MS in [9], we notice that there exist

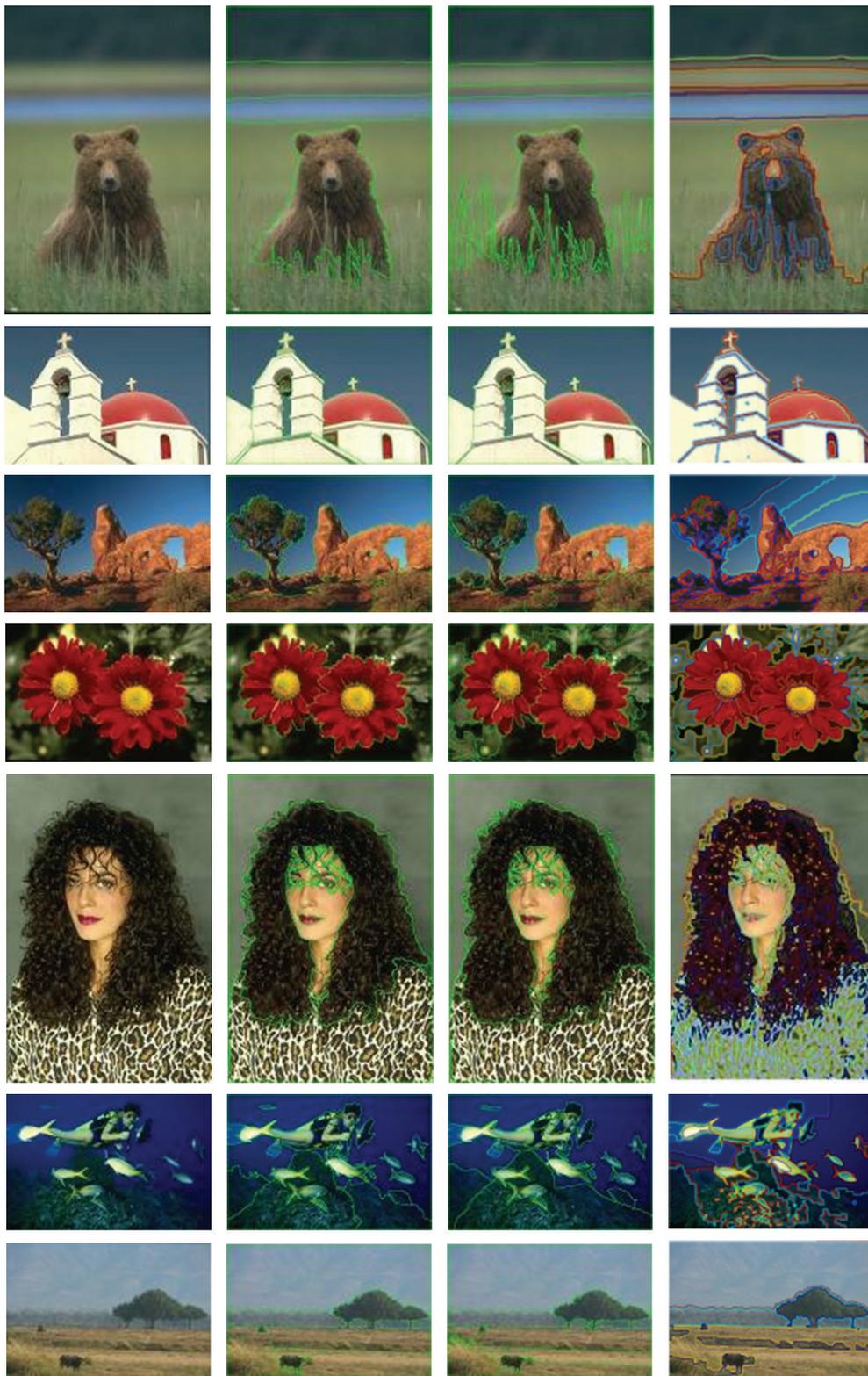


FIGURE 4: Segmentation results, comparing with human segmentations. Each row from left to right: original image, human segment, human segment, and our segment.

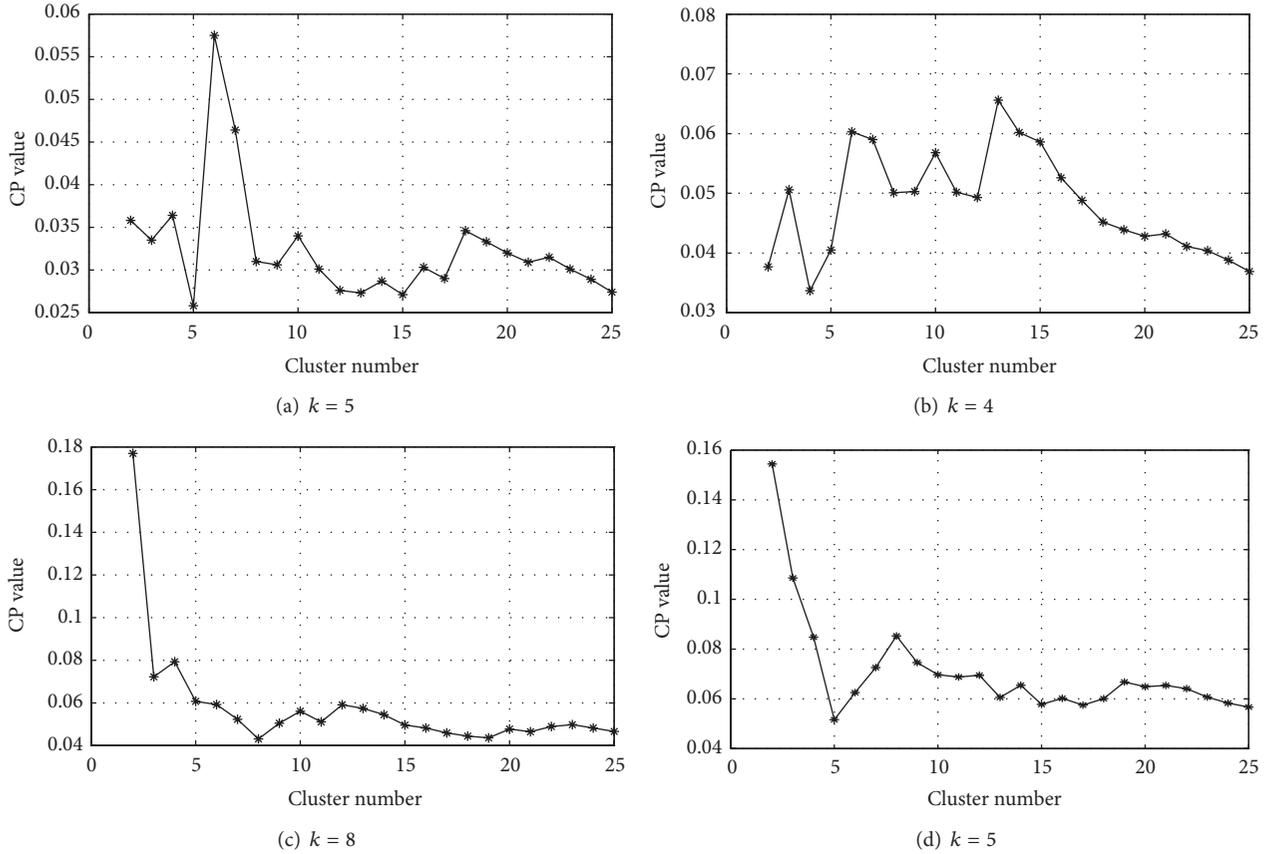


FIGURE 5: (a), (b), (c), and (d) are the cluster number determination of the images (a), (b), (c), and (d) of Figure 4, respectively. In each graph, x -axis represents the number of clusters, whereas y -axis shows the values of the compactness criterion, and $K_{\max} = 25$; the compactness criterion reaches the minimum value at the optimal cluster number. According to these results, the optimal cluster numbers are obtained as showed below each graph.

slightly unsmooth and serrated boundaries: see the MIT image in Figure 6 of [9], the room image in Figure 7 of [9] and the hand image in Figure 9 of [9]. Some errors are obvious in those segmentations. Notice that for both MS and the proposal, there exist over-segmentations due to a small gradient of illumination: see the table in Figure 7 of [9].

Second compare with CTM algorithm. Apart from the analysis in the introduction, the CTM algorithm solves the unsupervised problem by minimizing the lossy description length of the feature vectors, which will hold the performance of MDL. Except for the discussion on MDL based segmentation algorithm above, we notice that, for CTM algorithm, the experimentation shows that the results are sensitive to the parameters. Based on the segmentation results provided in [10], it can be seen that there are over-segmentations when $\gamma = 0.1$ and under-segmentations when $\gamma = 0.2$ (like the left picture of the second row in Figure 10 of [10]). For CTM algorithm, we observe that the accuracy of the final segmentation also depends on the accuracy of the boundaries and the regions of the pre-segments done before the merging process. Observing the boundaries of the church in the Figure 11 of [10], boundaries are unsmooth and incorrect. On the computation time, the proposed method clearly beats CTM algorithm without any sensitive parameter to tune.

4.3. Quantitative Evaluation. We now compare the quantitative performance of the proposed algorithm against MS and CTM. The comparison is based on four quantitative performance measures which have been used in many works as follows [10].

PRI (the probabilistic rand index). It counts the fraction of pairs of pixels whose labeling is consistent between the computed segmentation and the ground truth.

VoI (the variation of information). The metric defines the distance between two segmentations as the average conditional entropy of one segmentation given the other and thus roughly measures the amount of randomness in one segmentation which cannot be explained by the other.

GCE (the global consistency error). It measures the extent to which one segmentation can be viewed as a refinement of the other. Segmentations which are related in this manner are considered to be consistent, since they could represent the same image segmented at different scales.

BDE (the boundary displacement error). It measures the average displacement error of boundary pixels between two segmented images. Particularly, it defines the error of one

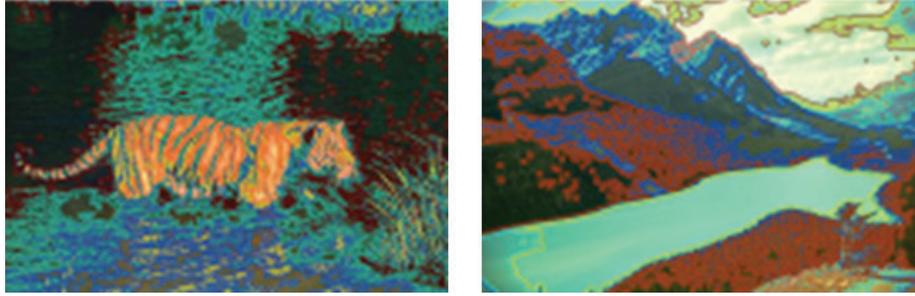


FIGURE 6: K -means clustering-based segmentations: unsmooth and discontinuous boundary and many outliers.



FIGURE 7: Compare K -means clustering-based segmentation (left) with KmsGC based segmentation (right).

boundary pixel as the distance between the pixel and the closest pixel in the other boundary image.

In the experiments, we set all parameters as in [10] or as the values recommended by the authors in their papers for the goal to achieve best results.

Table 1 shows the quantitative comparison of the proposed algorithm with MS and CTM algorithms on all the 300 images of Berkeley image database based on four indices, and in the table the bolded letters are the best value for each index. From the result we observe that the proposed algorithm KmsGC obtains competitive performances in terms of the four indices. In detail, the proposal obtains the best results with the PRI measure. On the BDE index, the proposal gets the best values, 0.5775 better than MS.

It is important to notice that the quantitative evaluation result also indicates that none of the algorithms is a clear winner in terms of all four indices. Notice that CTM algorithm shows different performance while the parameter γ is given different values. From the visually segmentation results in [10] we can see that the segmentation is coarser while γ is bigger. While $\gamma = 0.1$, CTM performs almost the same with MS algorithm on PRI, VoI, and GCE indices. But while $\gamma = 0.2$, CTM gets better performances on the first three indices but generates undersegmentation and loses accuracy

TABLE 1: Quantitative evaluation result of algorithms on the four indices (the best values for each index are in bold letters). PRI ranges between $[0, 1]$, the higher the better. VoI ranges between $[0, \infty)$, the lower the better. GCE ranges between $[0, 1]$, the lower the better. BDE ranges between $[0, \infty)$, in the unit of pixel, the lower the better.

Approach	PRI	VoI	GCE	BDE
Humans	0.8754	1.1040	0.0797	4.994
KmsGC	0.7712	2.5616	0.2932	9.1225
CTM ($\gamma = 0.1$)	0.7561	2.4640	0.1767	9.4211
CTM ($\gamma = 0.15$)	0.7627	2.2035	0.1846	9.4902
CTM ($\gamma = 0.2$)	0.7617	2.0236	0.1877	9.8962
MS	0.7550	2.4770	0.2598	9.7001

on BDE index. Overall the proposed KmsGC outperforms other algorithms on PRI and BDE indices.

We may look into the quantitative evaluation results and observe the relationship between the visual and the quantitative evaluation. According to the definition of the indices, PRI index measures the correctness of the pixel labeling, and VoI index measures the difference between two segmentations based on conditional entropy of the information. So a higher PRI value means more accurate segmentation results.

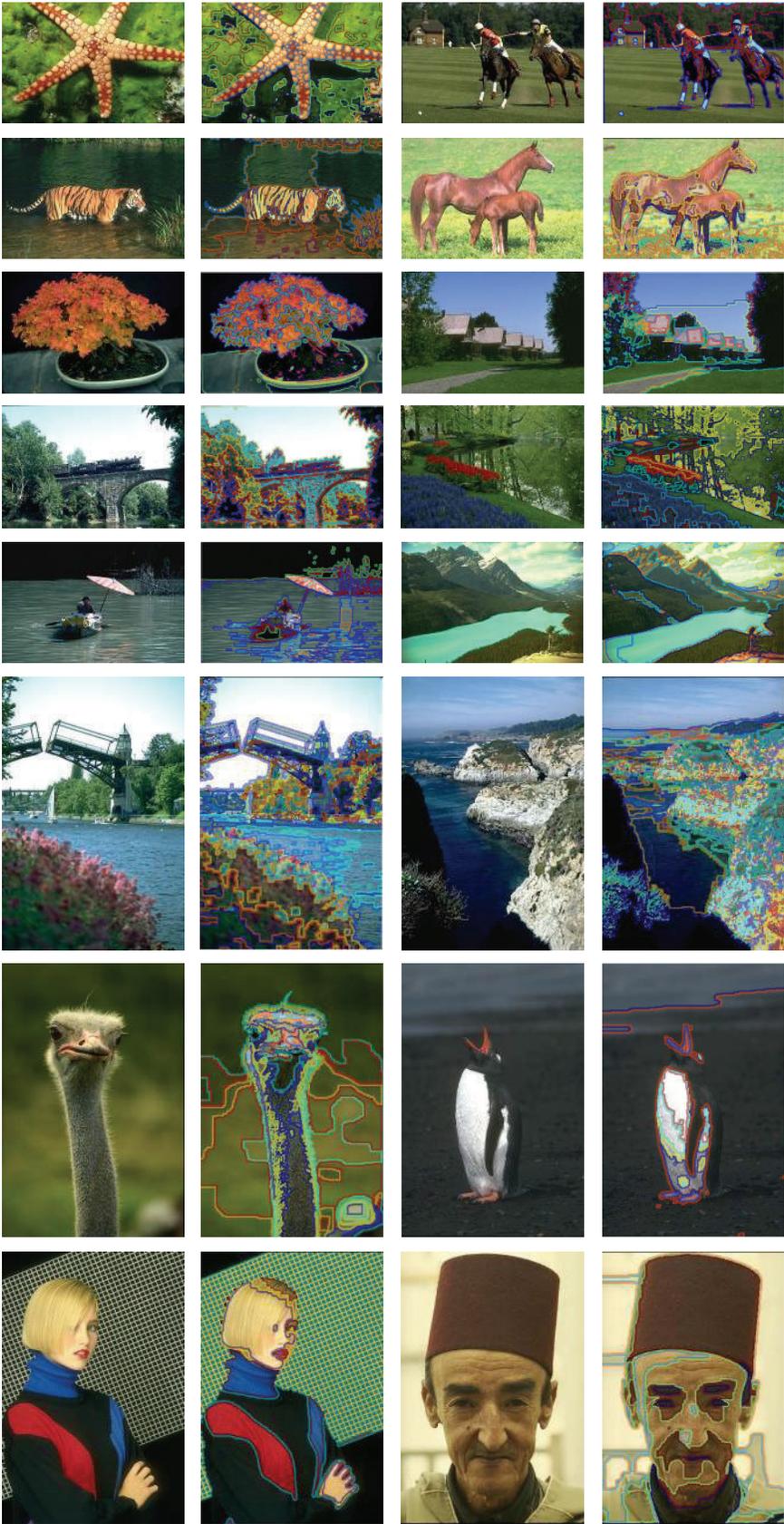


FIGURE 8: Examples of the segmentation results. 1st and 3rd columns are the original images; 2nd and 4th columns are the segmented images.

Benefiting from optimization of the information-theoretic criterion, CTM algorithm deserves the best performance on VoI index. BDE index measures the average displacement error of the boundary pixels. The proposal optimizes the boundary by an energy function, so the segmentation results demonstrate that the segmented boundaries match those of the human segmentations, which brings a best BDE value. As noticed in [10] that GCE index penalizes undersegmentation more heavily than oversegmentation and it does not penalize oversegmentation at all; that is, the highest score is achieved by assigning each pixel as an individual segment. Therefore, while we get a poor GCE, we obtain large segmented regions and smooth boundaries. Therefore, it may be concluded from the above analysis that the quantitative evaluation results are consistent with the visual evaluation results.

5. Discussion and Conclusion

We first discuss the conceptual advantages of the proposed method. The proposed KmsGC is a two-stage solution which combines K -means clustering technique with graph cut technique. In the first stage, a global and efficient initial pixel partition is obtained by K -means algorithm. In the second stage, the proposal casts the image segmentation constraints into the problem of energy optimization. The data penalty term expressing the constraints of region properties and the discontinuity penalty term expressing the boundary constraints are integrated in an energy function. The data penalty term is in charge of and experimentally produces compact clustering and accurate pixel label assignment, and the discontinuity penalty term is in charge of and experimentally brings accurate boundary location and smooth boundaries. The energy formulation is solved by a graph-based optimization method, graph cuts, to obtain a global minimum. A powerful α -expansion is run for the minimum cut of the graph and brings a fast segmentation.

For the unsupervised solution, a compactness criterion is constructed based on clustering principle to determine the optimal number of clusters, which is straightforward and ensures compact clustering. And K -means algorithm is applied to partition a cluster into two subclusters in the iteration, which is simple and efficient.

In general, the proposed model successfully incorporates K -means algorithm and graph cut technique into a new algorithm and hence owns the advantages of both of them, which provides fast, effective, and encouraging segmentation.

Comparing our segmentation results with the segmentation based on K -means algorithm will further demonstrate the advantages of the proposal. As K -means clustering algorithm partitions the pixels based on the features of single pixel, and it does not take into account the neighborhood relations among the pixels as well as the boundary constraints, it cannot produce smooth and continuous boundaries, and a lot of outliers will exist in the segmentation. Figures 6 and 7 show these defects. Comparing with the results in Figure 4, it is obvious that the proposed KmsGC solves these problems and provides satisfied segmentation.

A variety of experimental results demonstrate that the proposed algorithm KmsGC is effective and able to produce

desired image segmentation results. It is important to notice that we do not have any complicated parameter to tune, and the algorithm is straightforward and unsupervised.

However, some disadvantages exist in the proposal: (1) due to the random initialization of the K -means algorithm, the proposal could produce different segmentations in different running. To solve this problem, one can run the algorithm several times to obtain a good result; (2) as can be seen from the experimental result, a few outliers exist in some segmentations; (3) if a small gradient of illumination exists in a region of the image, the proposal is inclined to generate oversegmentation.

Oversegmentation is an intrinsic problem appearing in clustering-based segmentation. The technique of region merging has been brought up to solve this problem. In [37], the author merges adjacent regions with similar grey value level to reduce oversegmentation. However, merging the oversegmentations into an optimal segmentation is inherently difficult. In addition, fusing intensity, color, texture, and other features together to measure the similarity between the pixels can be a possible solution to this problem because more feature information will help to partition the image pixels better.

In this work, the human segmentations are assumed as the best perceptual references. To reach better performance, the segmentation result of the proposal has to be tuned to be closer to the human segmentation. So it is necessary to learn the patterns of human segmentation's behavior and then improve the proposal to get results in a more similar way to the humans. These are some of the challenging problems to be solved in the future.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors would like to thank the editors and the reviewers for their valuable reviews and comments. They would also like to deeply thank Dr. Dongquan Liu and Paul Liu for their great contributions to the improvement of this paper.

References

- [1] C. Petitjean and J. Dacher, "A review of segmentation methods in short axis cardiac MR images," *Medical Image Analysis*, vol. 15, no. 2, pp. 169–184, 2011.
- [2] H. A. Jalab and N. A. Abdullah, "Content-based image retrieval based on electromagnetism-like mechanism," *Mathematical Problems in Engineering*, vol. 2013, Article ID 782519, 10 pages, 2013.
- [3] J. Wang and M. F. Cohen, "Image and video matting: a survey," *Foundations and Trends in Computer Graphics and Vision*, vol. 3, no. 2, pp. 97–175, 2007.
- [4] J. Shotton, J. Winn, C. Rother, and A. Criminisi, "Textonboost for image understanding: multi-class object recognition and segmentation by jointly modeling texture, layout, and context,"

- International Journal of Computer Vision*, vol. 81, no. 1, pp. 2–23, 2009.
- [5] M. A. Ruzon and C. Tomasi, "Color edge detection with the compass operator," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '99)*, vol. 2, pp. 160–166, Fort Collins, Colo, USA, June 1999.
 - [6] S.-C. T. Cheng, "Region-growing approach to colour segmentation using 3-D clustering and relaxation labelling," *IEEE Proceedings: Vision, Image and Signal Processing*, vol. 150, no. 4, pp. 270–276, 2003.
 - [7] L. Vincent and P. Soille, "Watersheds in digital spaces: an efficient algorithm based on immersion simulations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 6, pp. 583–598, 1991.
 - [8] I. Vanhamel, I. Pratikakis, and H. Sahli, "Multiscale gradient watersheds of color images," *IEEE Transactions on Image Processing*, vol. 12, no. 6, pp. 617–626, 2003.
 - [9] D. Comaniciu and P. Meer, "Mean shift: a robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, 2002.
 - [10] A. Y. Yang, J. Wright, S. Sastry, and Y. Ma, "Unsupervised segmentation of natural images via lossy data compression," *Computer Vision and Image Understanding*, vol. 110, no. 2, pp. 212–225, 2008.
 - [11] P. Felzenszwalb and D. Huttenlocher, "Efficient graph-based image segmentation," *International Journal of Computer Vision*, vol. 59, no. 2, pp. 167–181, 2004.
 - [12] J. Shi and J. Malik, "Normalized cuts and image segmentation," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 731–737, San Juan, Puerto Rico, June 1997.
 - [13] Y. Deng and B. S. Manjunath, "Unsupervised segmentation of color-texture regions in images and video," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 8, pp. 800–810, 2001.
 - [14] Z. Yu, O. C. Au, R. Zou, W. Yu, and J. Tian, "An adaptive unsupervised approach toward pixel clustering and color image segmentation," *Pattern Recognition*, vol. 43, no. 5, pp. 1889–1906, 2010.
 - [15] G. P. Balasubramanian, E. Saber, V. Misis, E. Peskin, and M. Shaw, "Unsupervised color image segmentation using a dynamic color gradient thresholding algorithm," in *Human Vision and Electronic Imaging XIII*, B. E. Rogowitz and T. N. Pappas, Eds., vol. 6806 of *Proceedings of SPIE*, San Jose, Calif, USA, January 2008.
 - [16] Y. Boykov and M.-P. Jolly, "Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images," in *Proceedings of the 8th International Conference on Computer Vision (ICCV '01)*, vol. 1, pp. 105–112, Vancouver, Canada, July 2001.
 - [17] A. Delong, A. Osokin, H. N. Isack, and Y. Boykov, "Fast approximate energy minimization with label costs," *International Journal of Computer Vision*, vol. 96, no. 1, pp. 1–27, 2012.
 - [18] Y. Boykov and G. Funka-Lea, "Graph cuts and efficient N-D image segmentation," *International Journal of Computer Vision*, vol. 70, no. 2, pp. 109–131, 2006.
 - [19] C. Rother, V. Kolmogorov, and A. Blake, "GrabCut": interactive foreground extraction using iterated graph cuts," in *Proceedings of the 31st ACM International Conference on Computer Graphics and Interactive Techniques (ACM SIGGRAPH '04)*, pp. 309–314, Los Angeles, Calif, USA, August 2004.
 - [20] Y. Li, J. Sun, C. -K. Tang, and H. -Y. Shum, "Lazy snapping," *ACM Transactions on Graphics*, vol. 23, no. 3, pp. 303–308, 2004.
 - [21] J. Ning, L. Zhang, D. Zhang, and C. Wu, "Interactive image segmentation by maximal similarity based region merging," *Pattern Recognition*, vol. 43, no. 2, pp. 445–456, 2010.
 - [22] S. Vicente, V. Kolmogorov, and C. Rother, "Graph cut based image segmentation with connectivity priors," in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '08)*, pp. 1–8, Anchorage, Alaska, USA, June 2008.
 - [23] Y. Boykov, O. Veksler, and R. Zabih, "Efficient approximate energy minimization via graph cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 12, pp. 1222–1239, 2001.
 - [24] V. Kolmogorov and R. Zabih, "What energy functions can be minimized via graph cuts?" *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 2, pp. 147–159, 2004.
 - [25] A. Blake and A. Zisserman, *Visual Reconstruction*, MIT Press, Cambridge, Mass, USA, 1987.
 - [26] B. K. Horn and B. G. Schunck, "Determining optical flow," in *Techniques and Applications of Image Understanding*, vol. 0281 of *Proceedings of SPIE*, pp. 319–331, 1981.
 - [27] A. K. Jain, "Data clustering: 50 years beyond K-means," *Pattern Recognition Letters*, vol. 31, no. 8, pp. 651–666, 2010.
 - [28] M. H. Hansen and B. Yu, "Model selection and the principle of minimum description length," *Journal of the American Statistical Association*, vol. 96, no. 454, pp. 746–774, 2001.
 - [29] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, Morgan Kaufmann, 2006.
 - [30] S. Ray and R. H. Turi, "Determination of number of clusters in k-means clustering and application in colour image segmentation," 1999, <http://www.csse.monash.edu.au/~roset/papers/cal99.pdf>.
 - [31] Z. Volkovich, G.-W. Weber, R. Avros, and O. Yahalom, "On an adjacency cluster merit approach," *International Journal of Operational Research*, vol. 13, no. 3, pp. 239–255, 2012.
 - [32] Z. Volkovich, Z. Barzily, G.-W. Weber, D. Toledano-Kitai, and R. Avros, "Resampling approach for cluster model selection," *Machine Learning*, vol. 85, no. 1-2, pp. 209–248, 2011.
 - [33] J. Besag, "On the statistical analysis of dirty pictures," *Journal of the Royal Statistical Society*, vol. 48, no. 3, pp. 259–302, 1986.
 - [34] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721–741, 1984.
 - [35] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 9, pp. 1124–1137, 2004.
 - [36] S. C. Zhu and A. Yuille, "Region competition: unifying snakes, region growing, and bayes/mdl for multiband image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 9, pp. 884–900, 1996.
 - [37] L. Patino, "Fuzzy relations applied to minimize over segmentation in watershed algorithms," *Pattern Recognition Letters*, vol. 26, no. 6, pp. 819–828, 2005.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

