

## Research Article

# Self-Organizing Map-Based Color Image Segmentation with $k$ -Means Clustering and Saliency Map

Dongxiang Chi<sup>1,2</sup>

<sup>1</sup> School of Electronics and Information, Shanghai Dianji University, Shanghai 200240, China

<sup>2</sup> School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798

Correspondence should be addressed to Dongxiang Chi, chidongxiang@yahoo.com.cn

Received 24 February 2011; Accepted 24 March 2011

Academic Editors: Y.-S. Chen, M. Faundez-Zanuy, and S. Kwong

Copyright © 2011 Dongxiang Chi. 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.

Natural image segmentation is an important topic in digital image processing, and it could be solved by clustering methods. We present in this paper an SOM-based  $k$ -means method (SOM-K) and a further saliency map-enhanced SOM-K method (SOM-KS). In SOM-K, pixel features of intensity and  $L^*u^*v^*$  color space are trained with SOM and followed by a  $k$ -means method to cluster the prototype vectors, which are filtered with hits map. A variant of the proposed method, SOM-KS, adds a modified saliency map to improve the segmentation performance. Both SOM-K and SOM-KS segment the image with the guidance of an entropy evaluation index. Compared to SOM-K, SOM-KS makes a more precise segmentation in most cases by segmenting an image into a smaller number of regions. At the same time, the salient object of an image stands out, while other minor parts are restrained. The computational load of the proposed methods of SOM-K and SOM-KS are compared to J-image-based segmentation (JSEG) and  $k$ -means. Segmentation evaluations of SOM-K and SOM-KS with the entropy index are compared with JSEG and  $k$ -means. It is observed that SOM-K and SOM-KS, being an unsupervised method, can achieve better segmentation results with less computational load and no human intervention.

## 1. Introduction

Being a basis of high level image analysis tasks like object recognition, image retrieval, and scene understanding, and so forth, image segmentation is an important subject in image processing. Particularly, natural image segmentation is considered to be a difficult task, because there is not any *priori* knowledge in advance. It is found that in such cases salient objects are crucial for further image retrieval or scene understanding.

In an 8-bit RGB color space, each pixel in an image is represented by one of  $2^{24}$  combinations. Image segmentation can be considered to be a kind of clustering, which clusters similar pixels into same group. A successful segmentation depends on good selections of similarity measure, feature description of an image, evaluation of the segmentation, and *priori* information available. A natural image, not being a GIF animation image, magnetic resonance image, or any other man-made image, where advanced information is usually available, could be anything in the colorful, natural

world. It contains a large quantity of mixed elements with different color and light projected from a 3-dimension colorful world to a 2-dimension plane. The world is recovered from the compressed 2-dimension images although human are sometimes cheated by illusion tricks. When such images are submitted to a machine, there is always a tradeoff among segmentation results, computation complexity, and time cost no matter what method is selected.

Researchers have investigated natural image segmentation for many years. According to [1], segmentation methods are mainly classified as follows: feature-space-based techniques, image-domain-based techniques, and physics-based techniques. Clustering, histogram thresholding, and so forth are feature-space segmentation techniques; split-and-merge, region growing, edge-based and neural-network-based methods are image-domain segmentation techniques [1]. Some researchers categorized the segmentation methods based on two basic properties of the pixels in relation to their local neighborhood: discontinuity and similarity [2]. Corresponding methods based on discontinuity property of the

pixels are called boundary-based methods, while those based on similarity property are called region-based methods.

In fact, with the crossover of different disciplines, it is impossible to “segment” each segmentation technique clearly and each segmentation method can be classified into different categories by its different features. For example, neural-network method is, in fact, a kind of learning/training progress which touches every aspect of the segmentation, and it makes use of features or domain information of an image too. Methods involving multiple disciplines usually get better results. In recent years, methods like JSEG [3], mean shift [4], normalized cuts [5], and so forth have achieved certain success and are often used as benchmarks for natural image segmentation. Mean shift [4] detects modes of the probability density of colors occurring jointly in image and color domains and shows a good performance on segmenting the images with strong variations of density. Shi and Malik [5] treated the segmentation as a graph partitioning problem. A weighted undirected graph is constructed by taking the pixels as the nodes of graph. The graph is partitioned by optimizing the criterion of normalized cut based on the computation of eigenvalues. JSEG [3] makes a multiscale analysis in the image domain of a clustering map obtained from quantization in color domain. Consisting of two independent stages of color quantization and spatial segmentation, it is an unsupervised segmentation especially being robust when applied to scenes where texture predominates.

Our work in this paper is motivated on one hand by a performance improvement of natural image segmentation: a tradeoff between segmentation results, time complexity, and algorithm complexity. On the other hand, oversegmented or undersegmented image of the natural image slows down the further image retrieval or scene understanding. Operations of enhancing the salient objects and weakening the minor parts during segmentation are required to reduce the time of following processing.

This paper presents a clustering method SOM-K, which is based on self-organizing map (SOM) and  $k$ -means. Feature vectors of each pixel is first trained by an SOM neural network which clusters similar input feature vectors to be nearer with each other topologically, and then, the output vectors of SOM are clustered by  $k$ -means clustering method. Another proposed method with saliency map feature (SOM-KS) further improves the performance of SOM-K. Both SOM-K and SOM-KS are guided by an entropy index for image segmentation evaluation to select a best segmentation. To our knowledge, the combination of SOM and  $k$ -means was first appeared in [6]. A  $k$ -means clustering method with Davies-Bouldin (DB) validity index is implemented on SOM prototype vectors to segment remote sensing images and the results are declared acceptable and efficient. Although there were no further evaluation on the image segmentation results and the method was only applied on remote sensing image, their method was still appealing to our research. The  $k$ -means clustering method was modified and combined with an SOM to process  $L^*a^*b^*$  color space images [7]. It demonstrated some image segmentations and comparisons of different training parameters for SOM. In this paper, method SOM-K(S), which includes SOM-K and SOM-KS,

implements the SOM and  $k$ -means in a different way. It compromises the segmentation results and time cost with less complexity.

The main contributions of our proposed method are the following.

(1) SOM-K, a new unsupervised natural image segmentation method based on SOM and  $k$ -means. Intensity and  $L^*$ ,  $u^*$ , and  $v^*$  values of a color image are taken as features to be trained by a SOM network. The output prototype vectors are filtered by the hit map at first and clustered by the  $k$ -means method. A best image segmentation result can be obtained according to the entropy-based segmentation evaluation method. The method is proved to be robust to natural color image segmentation through experiments.

(2) A modified saliency map. The Itti-Koch visual attention model [8] is modified to be more efficient for our application. The Gaussian pyramid image is constructed via original image (level 0) from level 1 to level 5 and a group of orientation operators are applied to each level to get the orientation saliency map. We follow [8] using intensity, orientation, and color components  $R$ ,  $G$ ,  $B$ , and  $Y$  and adopt the contrast-based image attention model [9] within a  $3 \times 3$  window. After that, all saliency maps are resized to the original size and combined together to be a saliency map.

(3) SOM-KS, an unsupervised natural color segmentation method guided by image saliency map. Unlike other saliency map-based segmentation methods (refer to Section 2.2 for detail), the modified saliency map information is directly combined with intensity and  $L^*u^*v^*$  to segment a natural color image through SOM and  $k$ -means in an automatic manner. This method enhances the attractive objects in the image and restrains the less salient parts, which can reduce the processing workload for further processing.

The remainder of this paper is organized as follows. Section 2 presents related work. Section 3 describes in detail our proposed method. Section 4 analyzed the performance with comparisons with other methods. Section 5 is the conclusion.

## 2. Related Work

*2.1. Clustering and Image Segmentation.* Clustering methods provide us with a different view of the image segmentation. However, directly clustering methods like  $k$ -means and their variants are not acceptable considering the computational cost and *a priori* cluster number  $k$  needed. SOM with properties such as the input space approximation, topological ordering, and density matching, allied with the simplicity of the model and the ease of implementation of the learning algorithm, makes itself a promising clustering tool [10]. SOM is also helpful for visualization, cluster extraction, and data mining, and it has been proved to be successful for high dimensional data, where traditional methods may often fail or be insufficient.

Rarely simple SOM is implemented directly on image segmentation. Some researchers modified and expanded the typical SOM [11], while others combined SOM with other methods [6, 7, 10, 12–15]. Araújo and Costa [11] presented a new SOM with a variable topology for image

segmentation. The proposed fast convergent network is capable of color segmenting images satisfactorily with a self-adaptive topology. In [10], an SOM-based clustering method was applied to the spectral data of remotely sensed image. Accounting for the activity level (hits) information, the zero-hits nodes and heterogeneous nodes are filtered at first; secondly, the CD<sub>bw</sub> (composing density between and within clusters) clustering index is applied to the dendrogram to get the best cluster number. Although the target image segmented is simple comparing with natural images and some *priori* information could be obtained, their ideas have been enlightening to natural image segmentation research. Besides directly applying to color components, other used derived features in SOM [12, 14, 15]. Reference [12] presented an unsupervised image segmentation framework C<sub>Tex</sub> (based on color and texture), that is, based on the adaptive inclusion of color and texture in the process of data partition. In [14], a fully automatic three-level clustering algorithm for color-texture segmentation was presented. SOM is used to identify the number of components and initialize the Gaussian mixture model. Experimental results indicate that the proposed algorithm is efficient and competent with popular C<sub>Tex</sub> and JSEG algorithms.

Enlightened by [16], some SOM-based clustering methods focused on visualization of the input data by analyzing the SOM-derived parameters [17, 18]. Data topology is integrated into the visualization of the SOM, and thereby provides a more elaborate view of the cluster structure than existing schemes [17]. The prototypes are often combined with hit numbers to implements an automatic detection of clusters [18].

Recently, some researchers exploited color image segmentation with methods like distance metric learning [19], contour deformation and region-based method [20], morphological clustering [21], watershed variants [22, 23], edge information [24], local features measured by Gabor filter and clustered by expectation maximization (EM) [25], mean-shift variants [26], ant colony-fuzzy *c*-means hybrid [27], enhanced gradient network [28], dynamic region growth/multi resolution merging [29], and so forth broadened the road ahead.

**2.2. Saliency Map and Image Segmentation.** In most cases, the aim of image segmentation is object recognition, image retrieval, or scene understanding, and so forth, which serves as a necessary and the first step of high-level, object-based applications. Therefore, correct segmentation of salient objects in the image is more important than segmenting other minor parts correctly. It is an interesting topic covered by many researchers. Object recognition in an image follows top-down or bottom-up method. The first method needs *priori* knowledge of the top level, with face detector, human body detector, and so forth, to facilitate the recognition while the second method deals with natural image where no further information available.

Itti et al. introduced their bottom-up visual attention model [8] inspired by the behavior and the neuronal architecture of the early primate visual system. In their method, multi-scale image feature maps of color, intensity, and orien-

tation are extracted, and local spatial contrast is estimated for each feature at each location, providing a separate conspicuity map for each feature. These maps are combined to a single topographical saliency map that guided the attention focus in a bottom-up manner. However, the high computational cost and the parameter selection are still drawbacks of Itti-Koch model. After that, other derived image saliency map models were developed. In contrast to Itti-Koch attention model which derived attention based on the spatial location hypothesis, Sun and Fisher presented their mechanisms based on object-driven as well as feature driven. It is suggested that object-based and space-based attention can be integrated by using grouping-based salience to deal with dynamic visual tasks [30]. With advantage of not relying on either the parameters or rapidly salient objects detection, a spectral residual (SR) approach based on Fourier transform was proposed [31]. Hu et al. [32] led us to a different method to extract visual attentive regions in images using subspace estimation and analysis techniques. Ma and Zhang proposed a feasible and fast approach to attention area detection in images based on contrast analysis [9]. Some researchers also focused on saliency model of video [33]. Not trying to model human attention like traditional approaches, [34] proposed a new bottom-up validated stochastic model to estimate the probability that an image part is of interest. Reference [35] presented another new method focused on calculating the spatiotemporal saliency map of an image by its quaternion representation. Above all, some methods consider saliency over multi-scale image, while others over a single scale. In general, most methods use local contrast of image regions with their surroundings between multi-scale image using one or more of the features of color, intensity, and orientation.

As a basic step in image processing, image segmentation and object extraction are also facilitated by saliency map. On one hand, some authors defined their own saliency map for their research [36, 37]. On the other hand, some researchers exploited typical saliency map model, like Itti-Koch model [38–42] or spectral residual approach [43]. The saliency map being guidance to the image processing and object detection [44], many researches are expanded, and a lot of achievements gained. Applications like image retrieval [45–48], image retargeting [49], image content analysis [50], image fusion [51, 52], and image quality assessment [53] were all more or less based on the saliency map.

### 3. Proposed Method

In our proposed method, a color image features with/without saliency map are trained by an SOM neural network at first, the output prototype vectors are then filtered by a hits map, clustered by *k*-means accompanied with the guidance of entropy-based image segmentation evaluation index. Besides, the method also includes necessary preprocessing and postprocessing. The detail flowchart of the SOM-KS method is shown in Figure 1.

**3.1. Self-Organizing Map.** SOM, first put forward by Kohonen [54], is a kind of widely used unsupervised artificial

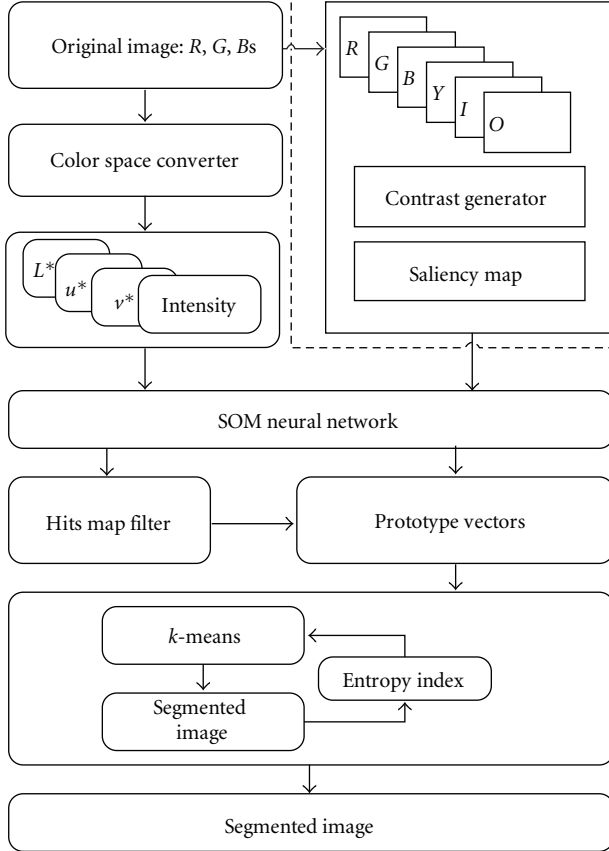


FIGURE 1: Flowchart of SOM-KS method. SOM-K follows the same flowchart except that the saliency map module is not included.

neural network. The map is a group of node units represented by prototype vectors lying in a 2-dimension space usually though occasionally nodes are set in one or multi-dimensional space. These units are connected to adjacent units by a neighborhood function. Prototype vectors are initialized with random or linear methods and “folded” in the 2-dimension space. Then, they are trained iteratively by randomly selected input samples sequentially or in batches and updated according to the neighbor function. After the training, prototype vectors become stable and “unfolded” themselves in the 2-dimension-space map. The typical features of SOM are topology visualization of the input patterns and representation of a large number of input patterns ( $p \times m$ , say, where  $p$  is the number of samples and  $m$  the dimension of the input pattern) with a small number of nodes ( $n \times m$ , where  $n$  is the number of nodes or prototype vectors). The most important attribute of SOM is that the input patterns which are similar in the input space are also nearby with each other topologically in the output space, the nodes map.

According to Kohonen [54], for each node  $i$ , there exists a prototype vector  $w_i = (w_{i1}, w_{i2}, \dots, w_{im})$ . For each input sample  $\mathbf{x}$ , a winner node,  $c$ , is chosen, using the similarity rule

$$c = \arg \min_i \{ \|\mathbf{x} - \mathbf{w}_i\| \}, \quad (1)$$

which means the same as

$$\|\mathbf{x} - \mathbf{w}_c\| = \min_i \|\mathbf{x} - \mathbf{w}_i\|, \quad (2)$$

where  $\|\cdot\|$  represents the Euclidean distance. The winner node  $c$  weight (prototype vector), together with the weights of neighbor nodes, is updated according to the following equation:

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + h_{ci}(t)[\mathbf{x}(t) - \mathbf{w}_i(t)], \quad (3)$$

where  $t$  indicates the iteration of the training process,  $\mathbf{x}(t)$  is the input sample of current iteration  $t$ , and the  $h_{ci}$  is the neighborhood function of the winner node  $c$ . The last term of (3) is a decreasing function of iteration time  $t$  and distance between the node  $i$  and the winner node  $c$ , the learning rate  $\alpha(t)$  and neighborhood function  $h$

$$h_{ci}(t) = \alpha(t)h(\|r_c - r_i\|, t), \quad (4)$$

where  $\alpha(t)$  is the learning rate,  $r_i$  and  $r_c$  are the positions of the node  $i$ , and the winner node  $c$  in the topological map, respectively.

Also, a more effective SOM can be trained in batches [54].

- (1) For the initial prototype vectors, take, for instance, the first  $K$  training examples, where  $K$  is the number of nodes (prototype vectors).
- (2) For each map unit  $i$ , collect a list of copies,  $N_i$ , of all those training samples  $x$  whose nearest prototype vector belongs to unit  $i$ .
- (3) Take for each new prototype vector the mean over the union of the lists in  $N_i$ .
- (4) Repeat from step 2 a few times.

This algorithm is particularly effective if the initial values of the prototype vectors are already roughly ordered even if they might not yet approximate the distribution of the input samples. It should be noticed that the above algorithm contains no learning rate parameter; therefore, it has no convergence problem and yields stable asymptotic values for the  $w_i$  than the original algorithm. Especially, a few iterations of this algorithm will usually suffice [54].

For every input sample, there exists one node with maximal output. The node is said to be hit by the input sample. The number of input samples hitting a node is named as the hits value of the node, which forms a hits map. The larger the hits value, the more input samples are represented by the node. According to the attribute of SOM, except those unhit nodes, every node of SOM represents a group of input feature patterns and the topology of the nodes in a 2-dimension SOM map shows the topology of a multi-dimensional space within input feature patterns, so the input feature patterns (pixels for image) can be clustered through clustering the nodes themselves. A common idea of the SOM prototype vectors and hit map is shown in Figure 2.

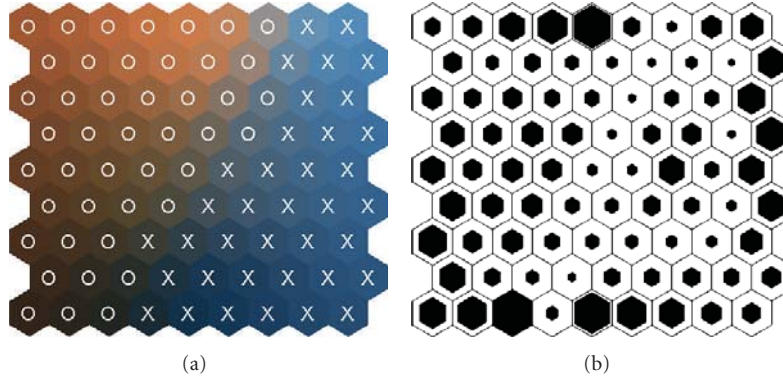


FIGURE 2: Clustered SOM nodes (a) and hit map (b) of image 295087. Color features  $R$ ,  $G$ , and  $B$  of the image 295087 (see Figure 6) are processed with the proposed SOM-K method. Pixels are clustered (or quantized) into 81 prototype vectors (nodes), whose colors (RGB in prototype vectors) are shown in (a). As can be seen in image 295087, there are mainly two colors with different tones, brown and blue in (a). Each prototype vector is a representative of a group of pixels. Total 81 representatives are further clustered with  $k$ -means into 2 clusters marked with  $\times$  and  $\circ$ . The right figure (b) shows the hits map of the SOM. The larger the black area in each node, the more it is hit by input patterns. The hits map is used as a filter to delete the prototype vectors. Those prototype vectors with zero hits are deleted, because they do not represent any input patterns.

3.2. *k*-Means and Image Evaluation. After the large quantities of pixel data are projected to a 2-dimension space to become a member in a group of nodes in a simple and fast way, a typical  $k$ -means method is adopted to cluster the prototype vectors. Clustering the SOM prototype vectors instead of directly clustering the data is a computationally effective approach [16].

As a kind of unsupervised learning method, clustering is divided to be hierarchical or partitional. The formal one can be agglomerative (bottom-up) or divisive (top-down). The latter one, partitional clustering, decides all clusters at once. Being a typical partitional clustering method,  $k$ -means method assigns each point to the cluster with the nearest center. The main steps of a standard  $k$ -means algorithm include [55] the following.

- (1) Set the number of cluster as  $k$  and randomly generate  $k$  clusters and determine the cluster centers, or directly generate  $k$  random points as cluster centers.
- (2) Assign each point to the nearest cluster center, usually calculated with Euclidean distance.
- (3) Re-calculate the new cluster centers.
- (4) Repeat steps (2) and (3) until convergence criterion is met.

The main advantages of  $k$ -means algorithm are its simplicity. Its disadvantages are heavy computation if the amount of data is large. It may not yield the same result with each run, since the resulting clusters depend on the initial random assignments. It minimizes intracluster variance, but it does not ensure that the result has a global minimum of variance. To overcome its disadvantages, the  $k$ -means is run for at least ten times to avoid instability caused by random assignments. Because the number of prototype vectors is very small, the computation cost is not a problem. With

the standard  $k$ -means algorithm, the prototype vectors are clustered from 2 clusters to  $\lceil \sqrt{n} \rceil$  ( $n$  is the number of SOM nodes) clusters, respectively, and image segmentation results with 2 clusters to  $\lceil \sqrt{n} \rceil$  clusters are obtained.

There are several metrics of cluster evaluation to decide the best cluster number, like DB index [56], and CDbw index [57]. DB index was most commonly used to estimate the best number of clusters in remotely sensed images clustering [6] or employed in combining with another index in the unsupervised classification method proposed [58]. As stated by [59], the DB index is suitable only for spherical clusters and is sensitive to noisy points. In practice, though, it is better to use the DB validity index values as a guideline rather than absolute truth [16, 60]. CDbw, on the other hand, puts emphasis on the geometric features of clusters, handling efficiently arbitrary-shaped clusters. This is achieved by representing each cluster by a certain fixed number of clusters rather than a single center point. It was declared to be a reliable index showing that it performs favorably in all cases selecting independently of clustering algorithm the scheme that best fits the data under consideration [57].

Besides, the results of image clustering can be evaluated directly by segmentation results with image segmentation evaluation index like entropy-based index [61] or quantitative-based index [62]. A lower quantitative value or entropy value leads to better segmentations. A typical quantitative image evaluation method is  $Q$  index defined as below:

$$Q(I) = \frac{1}{10000(N \times M)} \sqrt{R} \times \sum_{i=1}^R \left[ \frac{e_i^2}{1 + \log A_i} + \left( \frac{R(A_i)}{A_i} \right)^2 \right], \quad (5)$$

where  $N \times M$  is the size of image  $I$ ,  $R$  is the number of regions of the segmented image,  $A_i$  is the area (as measured

by the number of pixels) of the  $i$ th region, and  $e_i$  is the sum of the Euclidean distance between the RGB color vectors of the pixels of region  $i$  and the color vector attributed to region  $i$  in the segmented image. The  $R(A_i)$  represents the number of regions having an area equal to  $A_i$ . The  $Q$  index was declared to be an effective guide in tuning segmentation algorithms [62]. Compared to quantitative evaluation, the entropy method [61] is based on information theory instead of empirical analysis and was declared to be able to indicate local minima over a wider range of values than quantitative method, which cannot distinguish local minima when the number of regions is large. For the region  $j$  of a segmented image and the value  $m$  of the feature value in region  $j$ ,  $L_j(m)$  denotes the number of pixels in the region  $j$  that have a feature value  $m$  in the original image and  $V_j$  the set of all possible feature values in the region  $j$ .  $S_I$  and  $S_j$  denote the areas of image  $I$  and region  $j$ , respectively. The entropy of region  $j$  is defined as [61]

$$H(R_j) = - \sum_{m \in V_j} \frac{L_j(m)}{S_j} \log \frac{L_j(m)}{S_j}. \quad (6)$$

The expected region entropy of image  $I$  is

$$H_r(I) = \sum_{j=1}^N \left( \frac{S_j}{S_I} \right) H(R_j). \quad (7)$$

The layout entropy is defined as

$$H_l(I) = - \sum_{j=1}^N \frac{S_j}{S_I} \log \frac{S_j}{S_I}. \quad (8)$$

The entropy index of current segmentation is

$$E = w_l * H_l(I) + w_r * H_r(I). \quad (9)$$

The region entropy generally decreases with the number of regions, while layout entropy increases with the number of regions. Hence, two weights  $w_r$  and  $w_l$  can be used to counteract the effects of oversegmenting or undersegmenting [61]. Because we aim at salient objects in the images, the  $w_l$  and  $w_r$  are set 0.8 and 0.2, respectively, to obtain an undersegmentation. According to our experiments, the entropy-based index with different weights on  $H_r$  and  $H_l$  is found to generate fewer segmented regions than the quantitative method which tends to generate oversegmented regions. In fact, for most natural image segmentations in our experiments, the entropy-based segmentation nearly always leads to 2-cluster results with fewer regions, which makes sense in most natural images. As for DB index and CDbw index, they are fit for cluster evaluation instead of image evaluation. The entropy-based image evaluation index is adopted at last.

**3.3. Modified Saliency Map.** As discussed in Section 1, salient objects are crucial for further image retrieval or scene understanding. In many cases, an image is always oversegmented, and saliency map may provide valuable information. It can

be observed in Section 1 that for some specific applications, a modified saliency map is a good choice. In this paper, to enhance salient regions in a natural color image to avoid broken regions of an object, the typical Itti-Koch model is modified as followed.

At first, intensity image  $I$  is obtained through  $r$ ,  $g$ , and  $b$  components of an original image.  $I$  is calculated as

$$I = \frac{(r+g+b)}{3}. \quad (10)$$

Those pixels with intensity less than 10% of its maximum over the entire image have zero  $r$ ,  $g$ , and  $b$  value, because that hue variations are not perceivable at very low luminance (and hence not salient).

$R$ ,  $G$ ,  $B$ , and  $Y$  are formed in this way

$$R = r - \frac{(g+b)}{2},$$

$$G = g - \frac{(r+b)}{2},$$

$$B = b - \frac{(g+r)}{2},$$

$$Y = \begin{cases} \frac{(r+g)}{2} - \frac{|r-g|}{2} - b, & \text{if } \frac{(r+g)}{2} - \frac{|r-g|}{2} \\ & -b > 0, \\ 0, & \text{if } \frac{(r+g)}{2} - \frac{|r-g|}{2} \\ & -b \leq 0. \end{cases} \quad (11)$$

$Y$  is included here because there exists a so-called ‘‘color-double-opponent’’ system, where in the center of their receptive fields, neurons are excited by one color and inhibited by another color, while the converse is true in the surround. Such special and chromatic opponency exists for the red/green, green/red, blue/yellow, and yellow/blue [8]. According to our experiments, a saliency map with component  $Y$  and  $R$ ,  $G$ , and  $B$  shows good segmentation results than without it. Five groups of Gaussian pyramid level  $I(\sigma)$ ,  $R(\sigma)$ ,  $G(\sigma)$ ,  $B(\sigma)$ , and  $Y(\sigma)$  are generated from  $I$ ,  $R$ ,  $G$ ,  $B$ , and  $Y$ , respectively, where  $\sigma$  is Gaussian pyramid level index,  $\sigma \in [1, 2, 3, 4, 5]$ . Then, according to Ma and Zhang contrast-based saliency map [9], the center-neighborhood distance of  $I$ ,  $R$ ,  $G$ ,  $B$ , and  $Y$ , which include the luminant and chromatic saliency information, are calculated within a  $3 \times 3$  window by Euclidean distance.

As for orientations, a different approach of four line detection masks  $R_0$ ,  $R_{45}$ ,  $R_{90}$ , and  $R_{135}$  [63] is proposed to filter the intensity image  $I$  to simplify the calculation of Itti-Koch’s Gabor filters. Masks  $R_0$ ,  $R_{45}$ ,  $R_{90}$ , and  $R_{135}$  are

correspondent to four directions horizontal, 45°, vertical and 135°, respectively,

$$\begin{aligned}
 R0 &= \begin{bmatrix} -1 & -1 & -1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{bmatrix}, \\
 R45 &= \begin{bmatrix} -1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & -1 & -1 \end{bmatrix}, \\
 R90 &= \begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix}, \\
 R135 &= \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}.
 \end{aligned} \tag{12}$$

Each level of intensity image being filtered by four direction masks, contrast saliency map are calculated and combined to be an orientation saliency map  $O(\sigma)$ .

At last, all saliency maps are resized to the original image size (level 0) and combined together to be a saliency map. To enhance the high salient region and restrain the low salient region, the saliency map is square power transformed

$$S = \left( \frac{(I(\sigma) + R(\sigma) + G(\sigma) + B(\sigma) + Y(\sigma) + O(\sigma))}{6} \right)^2. \tag{13}$$

The proposed saliency map (flowchart shown in Figure 3) is proved through experiments to be efficient and effective in that at first, the luminance, chrominance, and orientation factors are taken into account to find the salient region; secondly, the simplified orientation calculation does not decrease the oriental salience; thirdly, the power transformation makes the salient region more significant without changing the smooth tone of the saliency map.

Figure 4 shows two natural images segmented by SOM-K and SOM-KS. Obviously, compared to SOM-K, SOM-KS with saliency map enhances the edge of salient regions and improves the segmentation performance.

**3.4. SOM-K and SOM-KS.** In the preprocessing stage, a median filter is applied to the RGB color space of an image separately with a  $3 \times 3$  window, pixel by pixel. The median filter erases salt-pepper noise and smoothes the image. It helps to decrease the instability of segmentation result caused by random noise.

There are a lot of color space to be selected as features in image segmentation, like RGB, CLE XYZ, CIE  $L^*u^*v^*$ , CIE  $L^*a^*b^*$ , YUV, and so forth. A color space is uniform, if equal distance in the color space corresponds to equal perceived color differences. Many color spaces are non-uniform. For example, RGB color space, displaying in screen instead of exhibiting the perceptual uniformity, does not model the way

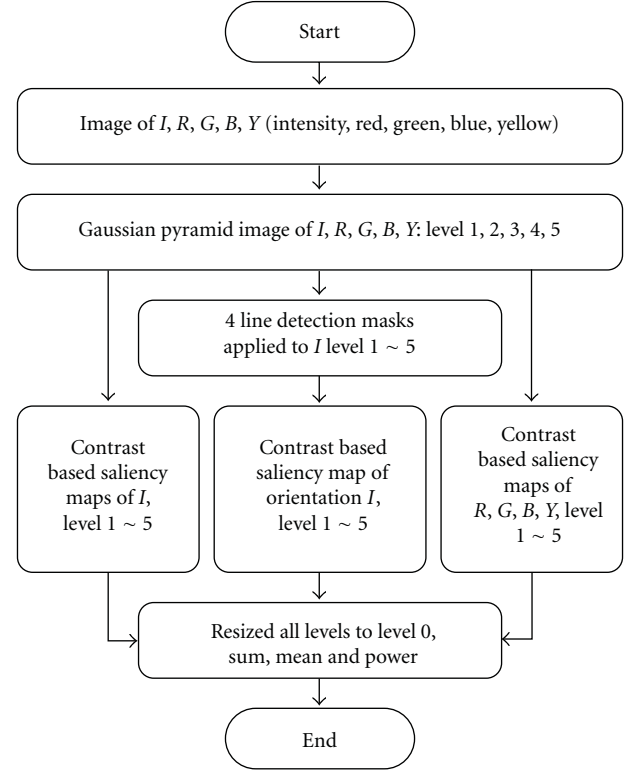


FIGURE 3: Flowchart of saliency map.

that human perceive colors. Comparatively,  $L^*u^*v^*$ , being a perceptual uniformly color space, performs better than other color spaces. According to our experiments, features  $L^*$ ,  $u^*$ , and  $v^*$  combined with intensity are a good choice. Figure 5 shows the comparison between different features applied to SOM-K and  $k$ -means methods of an image.

In this paper, RGB color space of the image is transformed into  $L^*u^*v^*$  color space. And then, the intensity and saliency map are integrated into color space to form input patterns  $T$  of SOM-KS method

$$T_{\text{SOM-KS}} = \{(i, l, u, v, s)_p\}, \quad p = 1, 2, \dots, P, \tag{14}$$

where  $P$  is the number of pixels in the image, and  $m = 5$ , the dimension of the input pattern including color space values  $i, l, u,$  and  $v$  and salience value  $s$ . The only difference between SOM-K with SOM-KS is that the feature  $s$  is not included in  $T$  of SOM-K

$$T_{\text{SOM-K}} = \{(i, l, u, v)_p\}, \quad p = 1, 2, \dots, P. \tag{15}$$

$T$  is trained with a 2-dimension hexagonal topology SOM with  $9 \times 9$  nodes and the output 81 prototype vectors are clustered with the  $k$ -means algorithm from cluster number 2 to 9. Blank nodes (not hit by any pixels) will be deleted before clustering. Last selection of the best cluster number of the image is entropy-based validity index [61]. The entropy-based metrics helps to find the best segmentation. In most cases, it selects 2 as best clusternumber. Note that cluster

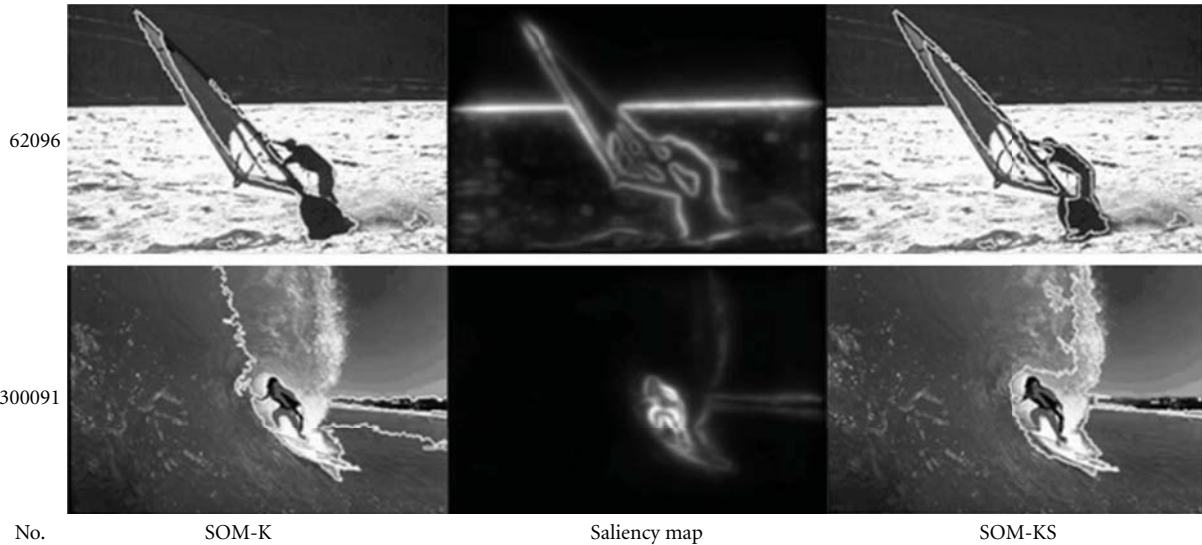
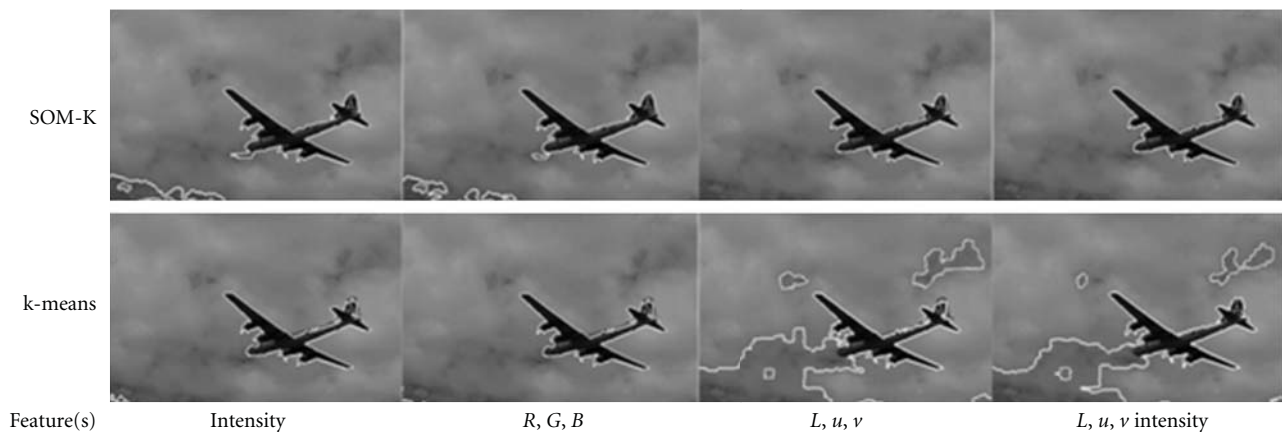


FIGURE 4: Enhancements made by saliency map.

FIGURE 5: Comparison between different features applied to SOM-K and  $k$ -means.

number does not necessarily mean region number in an image. The image may be divided into many regions though any of them is attributed to one of two clusters. Two clusters imply two regions at least. At the same time, three or more clusters do not ensure more regions than region number of two clusters. At the end of segmentation, those regions with area less than 0.1% of the original image area are neglected as a postprocessing step.

#### 4. Performance of Color Image Segmentation and Comparisons

**4.1. Segmentation Experiments and Analysis.** Our experiments take natural color images of Berkeley segmentation dataset (BSD) [64] as source images. The goal of the dataset and benchmark is to provide an empirical basis for research on image segmentation and boundary detection. The dataset images are divided into a training set of 200 images and a test set of 100 images [64]. All 100 test images are processed

with four methods: the proposed method SOM-K(S), SOM-KS is with saliency map as features while SOM-K not, JSEG method, and  $k$ -means method. Methods SOM-K, SOM-KS, and  $k$ -means are implemented based on MATLAB 6.5 and SOM toolbox 2.0 [65], while JSEG is based on window console. For JSEG, being a widely used image segmentation benchmark, images processed with default parameters are found to be oversegmented. So, according to their documents, the color quantization threshold and number of scales are set to 600 and 1, respectively, to obtain coarse segmentation images except that 3 out of 100 test images are not segmented by this setting (including 196073: sand snake, 38082: reindeer grassland, 227092: jar). The clustering number of the  $k$ -means method varies from 2 to  $\lceil \sqrt{n} \rceil$  clusters and adopts an entropy-based image segmentation evaluation index, just as in the proposed SOM-K(S) method. The main parameters setting of different methods are listed in Table 1. In most cases, the segmented images always tend to have less regions corresponding to 2 clusters according to



entropy-based validity index, either for  $k$ -means or SOM-K(S) methods, as discussed before.

The instability of the  $k$ -means method is overcome by running for at least ten times, either for SOM-KS, SOM-K, or for  $k$ -means. It is found that for most images, the segmentation will “converge” to one or several results for the proposed methods.

Because an evaluation of image segmentation is still more subjective than objective, the segmentation results are roughly categorized into three types: good, common, and unacceptable by human selection. The “converged” results of three categories are human selected with number of 40, 35, and 25, respectively. Figures 6, 7, and 8 shows some segmentation results of three categories.

As can be seen in Figure 6, the image segmentations with SOM-K(S) show better performance compared to JSEG. In fact, JSEG has consistent segmentation performances throughout 100 test images. Usually, it shows oversegmentation, but the shape of main objects in the image can be recognized clearly. The  $k$ -means shows similar segmentation performance as compared with the SOM-K(S) in many images although the time cost is higher. The entropy evaluation index involved, the segmentation results of SOM-KS tend to fewer regions compared to JSEG in most cases. Compared to SOM-K, SOM-KS plays better with the saliency map in following aspects (in Figure 6).

- (1) With the help of saliency map, the SOM-KS segments the images more precise in detail and object contour is enhanced to be complete and perfect. For example, in image 3096, a zigzag along the airplane contour is modified clearly in SOM-KS, compared to other three methods. In 69020 and 160068, it can be seen that the unconnected contours of kangaroo and big cat segmented by other three methods are completed by SOM-KS. So does images 62096, 300091, and so forth. With saliency map, the main object in the image is segmented clearly.
- (2) With the help of saliency map, nonsalient regions are restrained. This leads to a crisp and clean results compared to other three results, like in 14037, 58060, and 167083.
- (3) With the help of saliency map, the main object in the image stands out from the background. It can be found that in Figure 6, most of images either have a main object or the image foreground is clearly different from the background. This characteristic is compatible with the SOM-K(S) which clusters the pixels with SOM at first. A clear background and foreground facilitates the SOM to cluster.

Images in Figure 7 are complex than those in Figure 6. It is observed that the segmentation results are divergent within or between different images. For example, in images 38092, 41069, 123074, 175043, and 241048, some parts are segmented while others not. Compares to JSEG, the results are not satisfied though the SOM-K(S) keeps its characteristics, like fewer objects, better contour still. In

41069, for example, a little creature does not appear in SOM-K with a good segmentation of the rock but stands out in SOM-KS, and the rock is not segmented correctly. This case also happens to other images in Figure 7. It is difficult to decide whether the segmentation is good or not. In one hand, it surely clarifies some important objects in the image; on the other hand, it misses some other region boundaries. It also happens to JSEG in some images.

In Figure 8, the segmentation results are unacceptable for all methods. They do not tell useful information through the segmented image. The most important thing is that all these methods are incomparable to each other via the segmentation results themselves. These images bear the stamp of complex background, foreground, and the similar color, luminance. Some images are even difficult to be segmented by human, let alone machine itself without any *priori* information or help from human interaction. JSEG keeps its performance in this part and several segmented images are recognizable like 69015 and 108082.

**4.2. Entropy Index Comparison.** No matter for good or for unacceptable segmentation, comparisons among them are subjective. Objective comparisons are made with entropy index [61] in this part. Though index itself is not comparable to human selection, it is a kind of objective measure after all. Figure 9 shows the entropy index comparisons of JSEG (red circle),  $k$ -means (magenta square), SOM-K (green downward-pointing triangle), and SOM-KS (blue upward-pointing triangle) methods. The entropy indexes of 100 segmented images are computed with the original images and their labeled maps. For entropy index, a small value means a better segmentation. Apparently, for most images, the SOM-K, SOM-KS, and  $k$ -means show good and stable performances with low values than JSEG does. JSEG shows a scattered results ranging from 0.8 to 2.1 compared with the other 3 methods, which obtain a centralized results ranging from 0.9 to 1.6. The reason for this is that the  $H_I$  value of the entropy index with weight 0.8 is 0 for 1 region (unsuccessful segmentation with JSEG like 227092 and 196073, 32nd and 58th image in Figure 9) and a higher value for more regions. The  $H_r$  value with weight 0.2 varies with different images and different region. Smooth region means low  $H_r$  value (167062, 28th image in Figure 9), while region with more details leads to high  $H_r$  value. That is, the region number (with high weight) and region attributes (with low weight) have impact on the entropy index value. For JSEG method, the region numbers of 100 images range from 1 to 8. For the other 3 method guided by the entropy index, nearly all cluster numbers are 2, which means that most region number is 2 or a little more. So, most entropy index values of these 3 methods are small and centralized in a narrow range compared with JSEG method. A close look at the right figure shows that the SOM-K, SOM-KS, and  $k$ -means have similar entropy index values, and in many cases, they are superimposed with each other. This also agrees with their segmentation results. Also, we can see from the right figure that most SOM-KS entropy index values are a little larger than SOM-K and  $k$ -means. This is because that the segmentation results of SOM-KS are usually including an

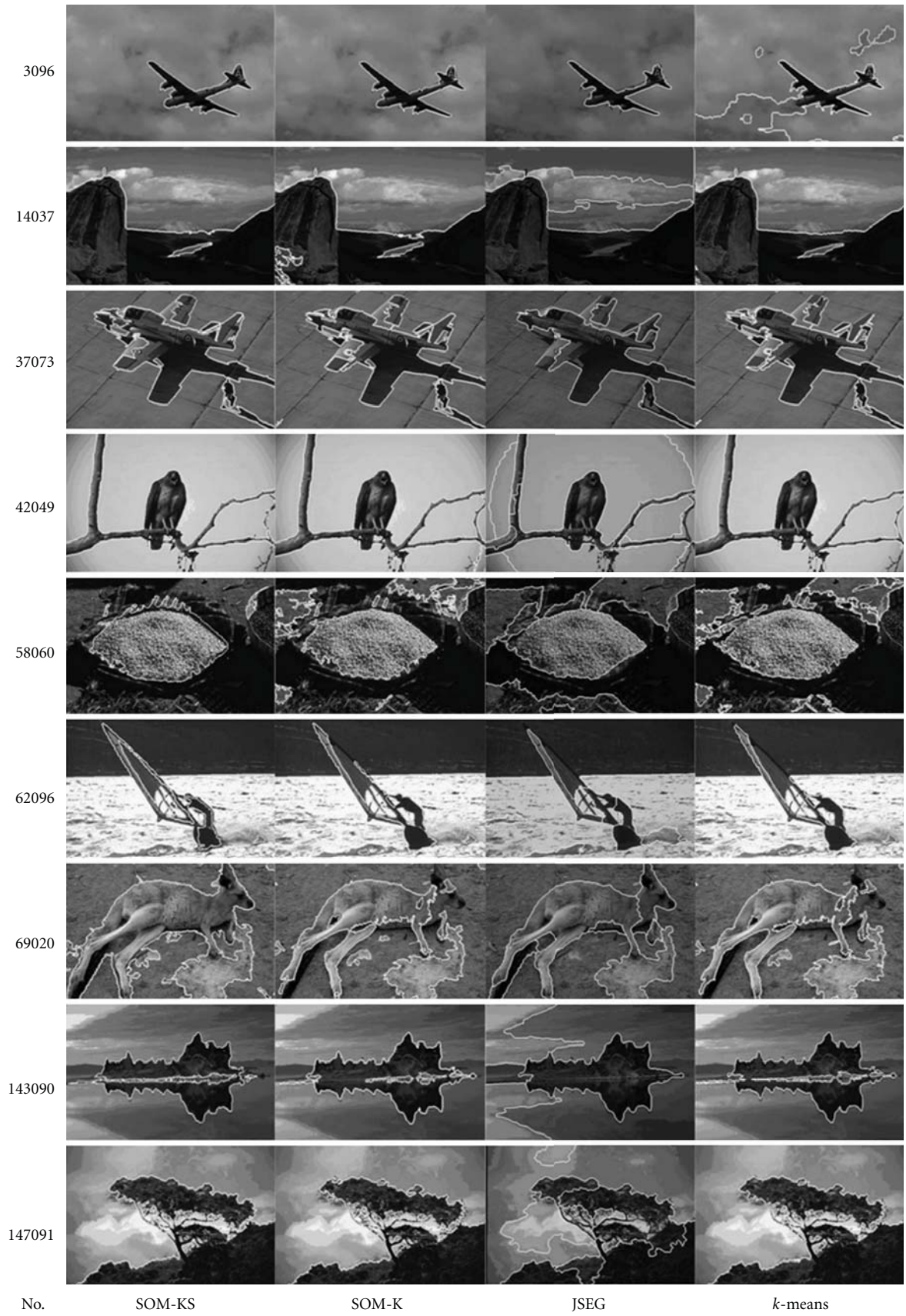


FIGURE 6: Continued.

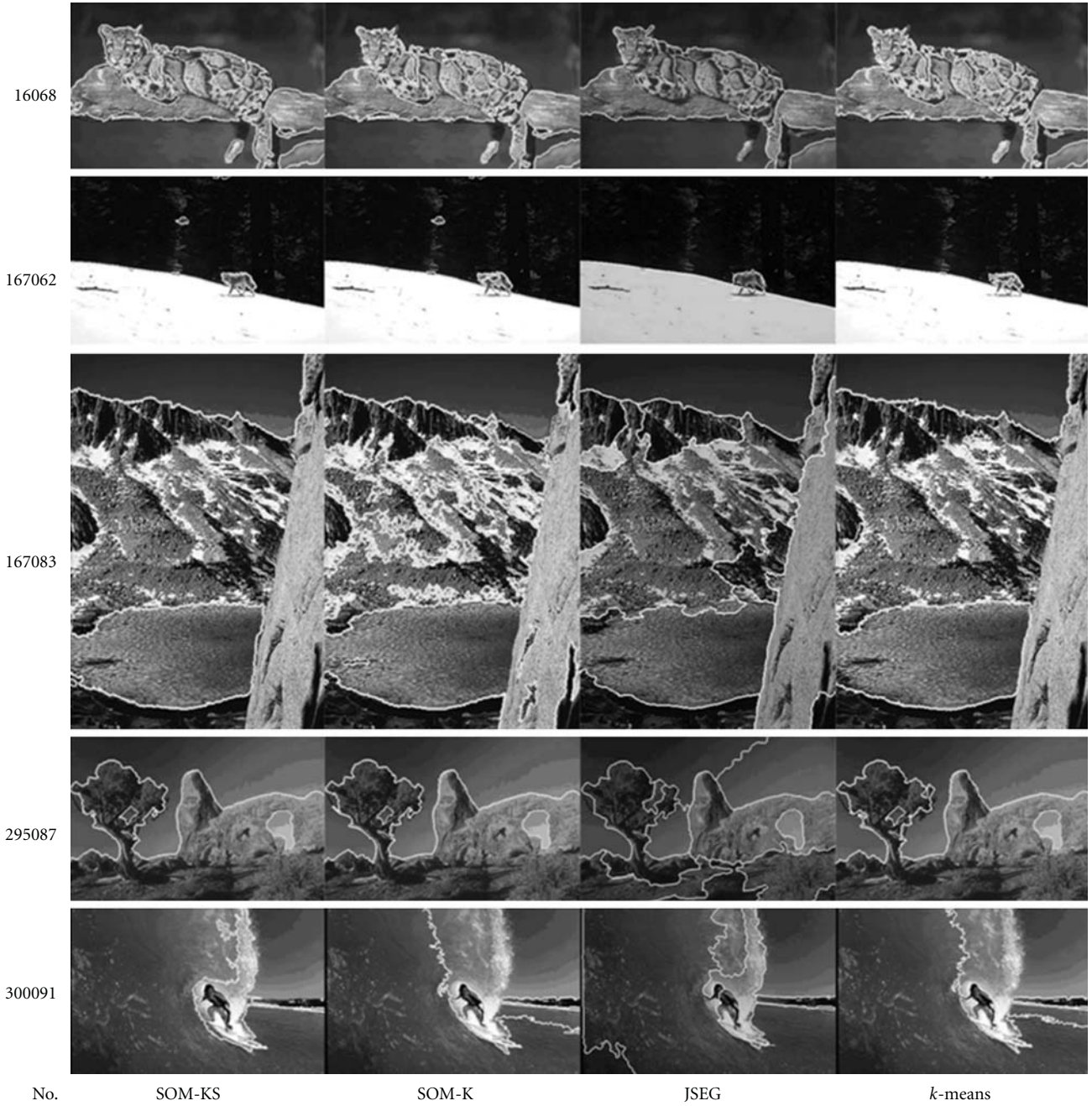


FIGURE 6: Good segmentation results.

TABLE 1: Parameters setting of different methods.

method	SOM-K(S)	k-means	JSEG
parameters setting	Preprocessing: normalize the variance of the variable to unity and its mean to zero. Topology: $9 \times 9$ nodes, hexagonal topology. Initialization and training: linear initialization, batch training. Training epoch: rough training 4, fine tuning 2, k-means: cluster number range: 2 ~ 9, squared Euclidean distance, online update phase and batch update phase, initialize with selecting $k$ observations from dataset at random.	Preprocessing: normalize the variance of the variable to unity and its mean to zero. Cluster number range: 2 ~ 9, squared Euclidean distance, online update phase and batch update phase, initialize with selecting $k$ observations from dataset at random.	default setting except $q = 600$ , $l = 1$ .

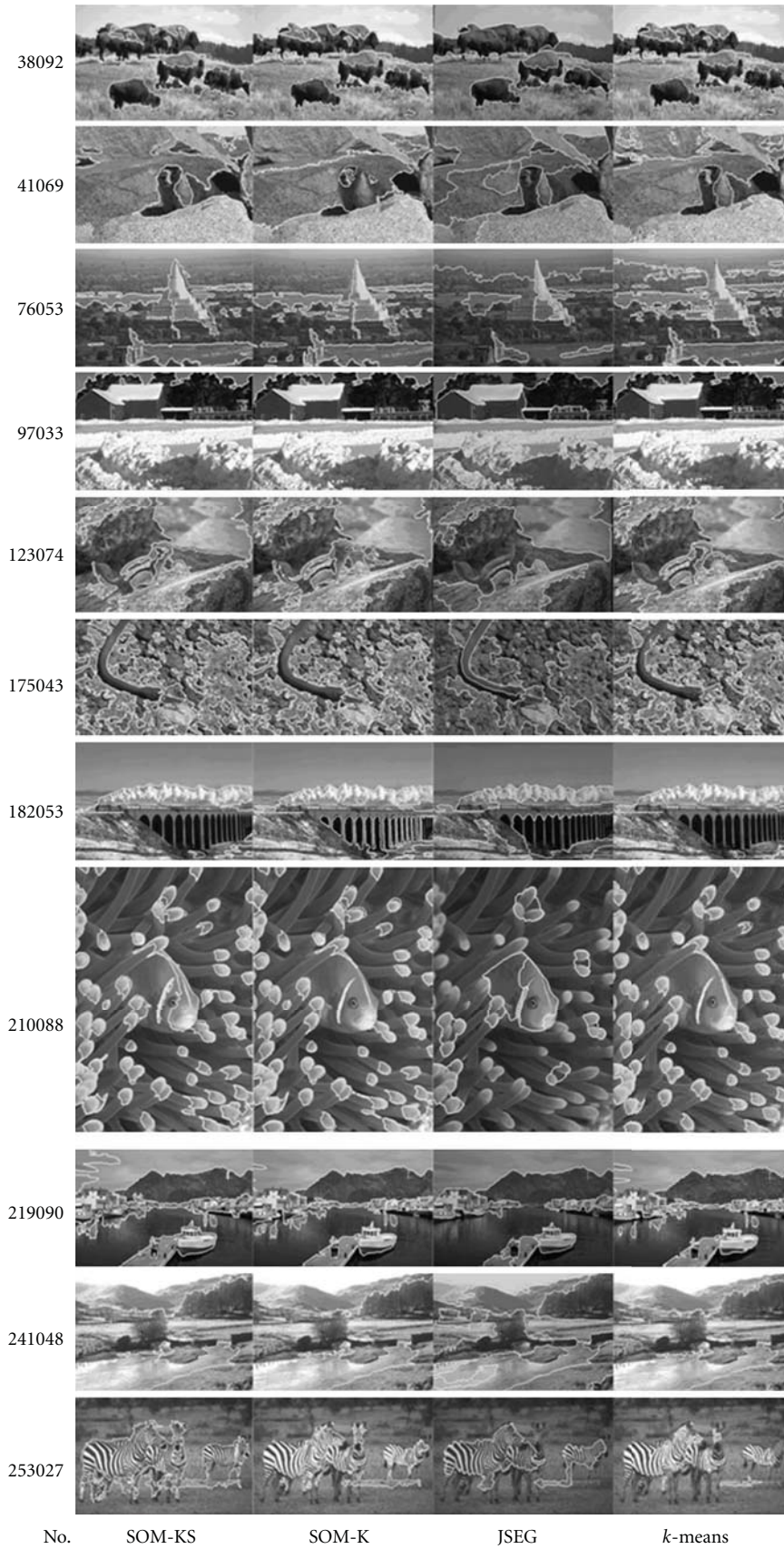


FIGURE 7: Common segmentation results.

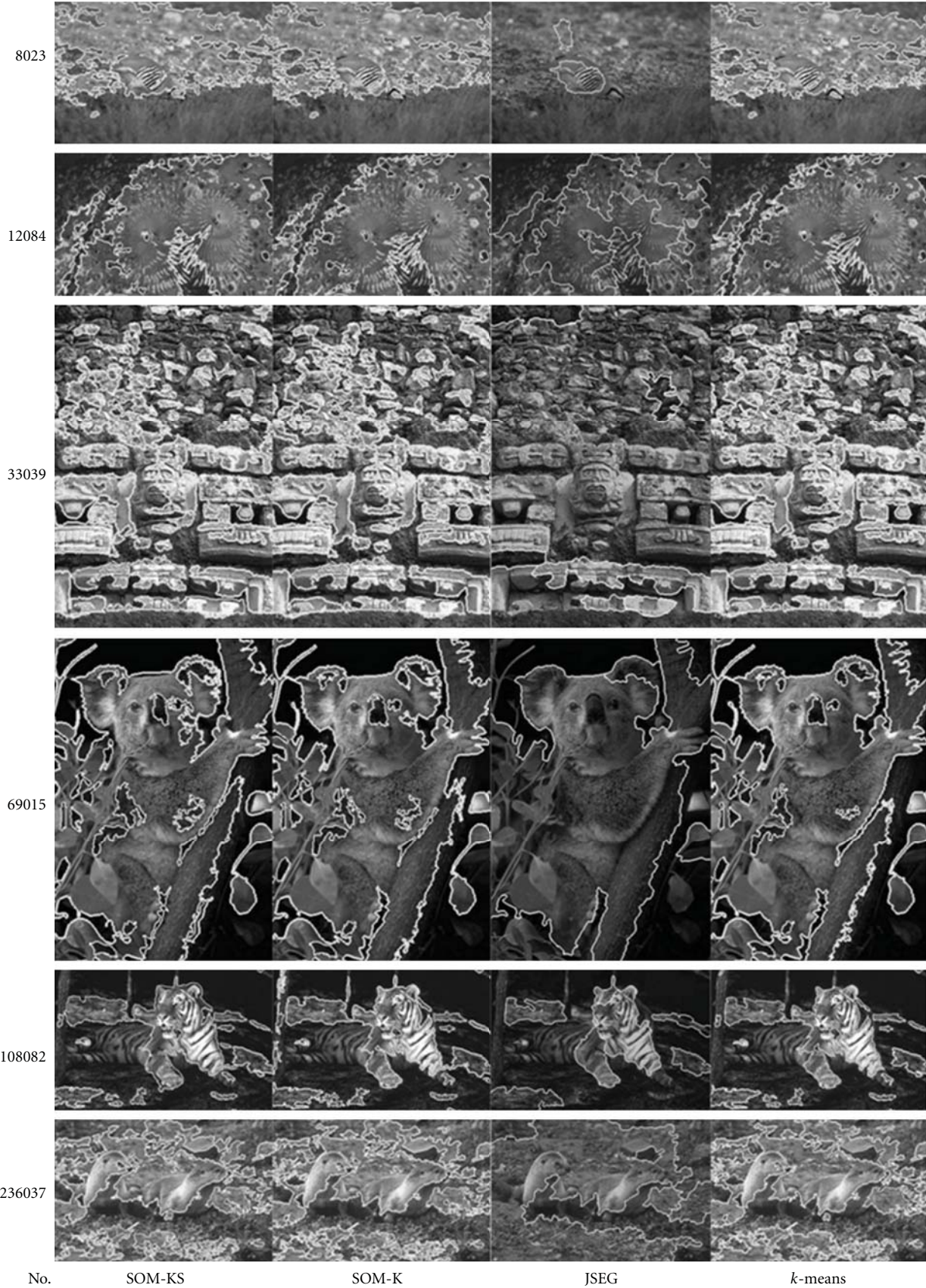


FIGURE 8: Unacceptable segmentation results.

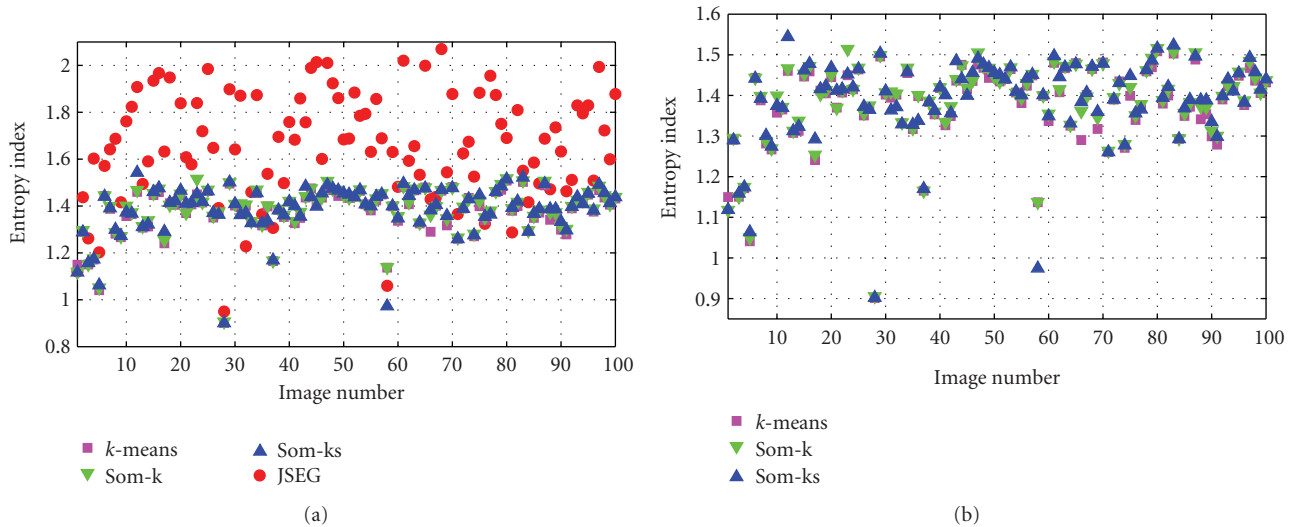


FIGURE 9: Entropy index comparisons of JSEG,  $k$ -means, SOM-K, and SOM-KS (a) and a close look of indexes without JSEG (b).

TABLE 2: Average computation time of different methods on 100 test images.

Method	SOM-KS	SOM-K	$k$ -means	JSEG
Time (seconds)	16.1	7.8	33.7	<8.4

object and a background. This leads to more variances within background and object itself that is,  $H_r$  values are a little larger than those segmentations of SOM-K and  $k$ -means without that much variances within each regions. After all, the so-called objective entropy index is still far away from human selection. Every image has its own attributes which have complex influences on the entropy index.

**4.3. Computational Complexity.** Computational complexity is a key characteristic in assessing a method. The SOM-KS, SOM-K, and  $k$ -means run in MATLAB platform and the JSEG is under a window console environment. The computer is with Microsoft Window-XP Professional 2002 operating system and Intel Duo CPU P8700@ 2.53 GHz, 2.89 GB Memory. All 100 test images are tested and the average computation time of these methods are listed in Table 2. All test images of BSD have same area, 154401 ( $481 \times 321$ ) pixels.

Computational complexity of  $k$ -means is  $O(pm \sum_{k=2}^C k)$ , where  $p \times m$  is dimension of samples,  $k$  is possible trials from 2 to  $C$ . For SOM-K combining SOM with  $k$ -means, the computational complexity is  $O(pmne)$ , where  $n$  and  $e$  are numbers of nodes and training epochs. That is, either  $k$ -means or SOM-based methods have linear computational complexities. In practice, the running time varies due to implementations and different parameter settings of each algorithm. Cores of both SOM and  $k$ -means are implemented with MATLAB and tuned carefully. As can be seen in Table 2, JSEG keeps a relatively high time-performance-ratio compared to other methods though its segmentation

results need further processing. SOM-K needs a little less time than JSEG, because SOM simplifies and accelerates the clustering process. Even with  $k$ -means clustering at second step, a small number of prototype vectors do not cost too much time. On the contrary,  $k$ -means alone costs more time than other methods. Compared to SOM-K, SOM-KS needs more than double times to finish segmentation. Apparently, the excess time comes from the saliency map. That is, a cost for a precise segmentation. It deserves a good result with time cost. In summary, we can see from the Table 2 that compared with  $k$ -means, SOM-K has similar segmentation performance and saves more time than  $k$ -means does. SOM-KS further enhances the segmentation performance with moderate increase in computational time and still being low than  $k$ -means. SOM-K uses similar time with JSEG, but with substantial segmentation improvement.

**4.4. Limitations.** As discussed before, the SOM-K(S) method can deal with most of the natural color image in the test dataset. But the involved  $k$ -means method leads to unexpected results. This comes from the random initialization of  $k$ -means. Figure 10 shows different segmentation results of same images, with same SOM-K(S) and different initialization of  $k$ -means. Occasionally, the results are confusing because on one side, the image could be right segmented but on the other side, it could be wrong. For example, in row 3 of Figure 10, the left circle land can be a correct segmentation if it is taken as a whole; also, the right circle land can be correct if you are serious about the detail of the circle land. A similar case appears in row 4 and 5, the person, jar, and red fish. If you focus on the person, the 1st and the 3rd are right segmentation. On the contrary, if you focus on the background, the other two may be a good segmentation. So does other images and vice versa. Strictly speaking, segmentations in Figure 10 are not good, but it depends on what people are interested in. If they are interested in the

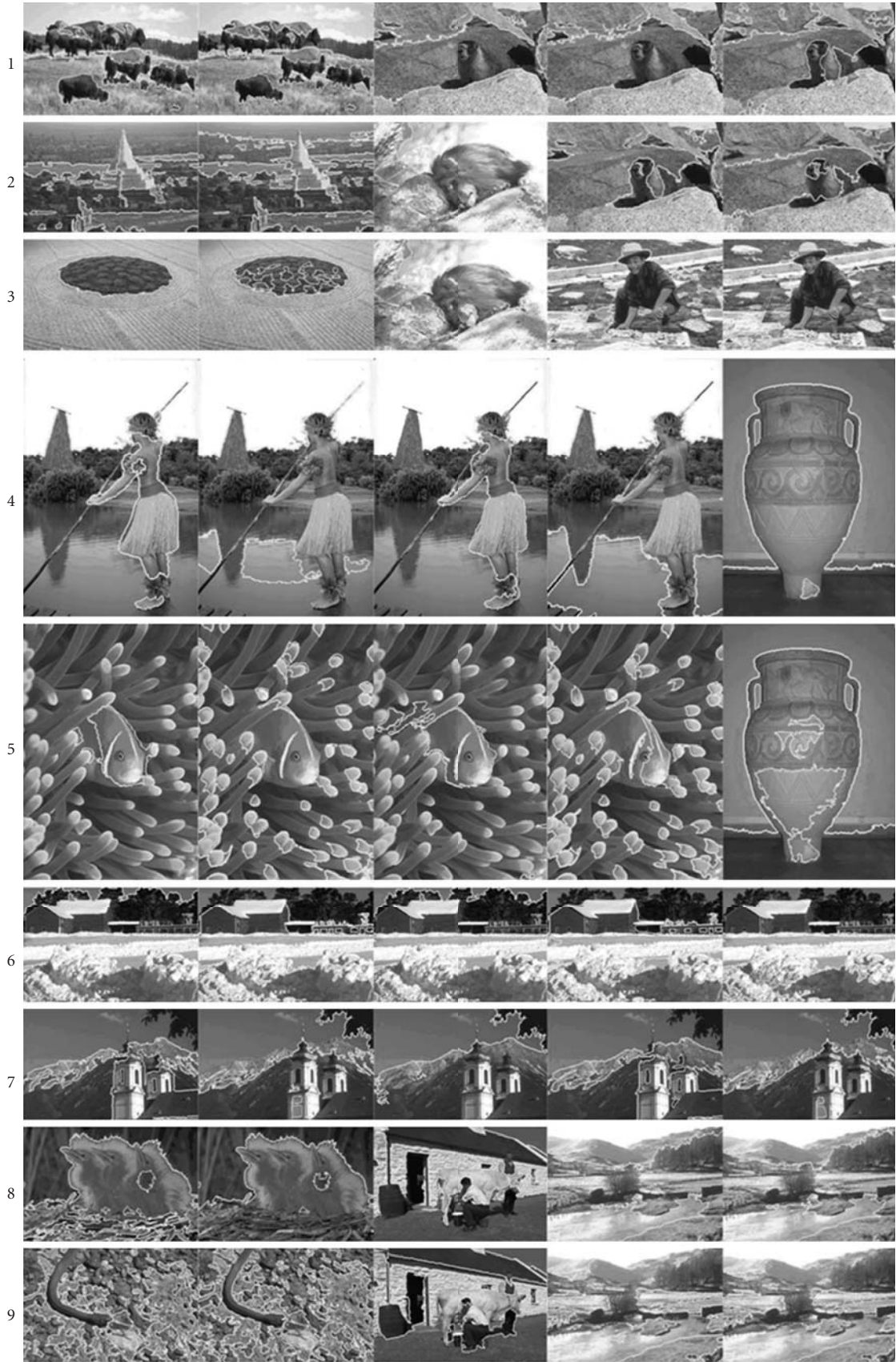


FIGURE 10: Different segmentation results of SOM-K(S).

objects in images, they might be good. If people are interested in the segmentation in whole, they are not.

As mention in [12, 66], color information are not enough to segment complicated natural scenes. The features of intensity and  $L^*u^*v^*$  in the proposed SOM-K(S) is not enough. So, a further description of a natural image with texture, and so forth, can improve the segmentation result. The idea in [11] with an optimum self-adaptive topology is also enlightening.

## 5. Conclusion

In this paper, a natural color image segmentation method based on SOM and  $k$ -means clustering is proposed. The method trains features of intensity and  $L^*u^*v^*$  color space with SOM neural network, and the output prototype vectors are clustered with  $k$ -means method (SOM-K). A variant of the proposed method is to combine a modified saliency map with intensity and  $L^*u^*v^*$  color space as new features to enhance the segmentation (SOM-KS). SOM-K(S) method is proved to be effective by experiments on the BSD. The saliency map enhances the segmentation results with more precise segmentation, and at the same time, the salient object in the image stands out and other minor parts are restrained; that is, over-segmentation is reduced in such areas. Segmentation evaluations of entropy index are compared with JSEG, SOM-K, SOM-KS, and  $k$ -means. The computational complexity is measured by the computation time and compared with each other. It is shown that the proposed SOM-K(S) method, being an automatic method, gets better segmentation results with less time needed and no need to set any parameters in advance. The only limitation is that it relies heavily on the initialization of  $k$ -means. It takes several times to obtain a better result. More segmentation images and source code of the proposed method are available on the web at <http://sites.google.com/site/chidongxiang/>.

## Acknowledgments

The research was supported by the Science and Research Program of Shanghai Dianji University under Grant no. 07C402 and Shanghai Dianji University Key Discipline Program no. 10XKF01. The authors thank Professor Xue Ping of Nanyang Technological University for his guidance and kindly help and thank Luo Ye, Wang Junyan, Huang Wei, and Shen Minmin of Nanyang Technological University for their valuable suggestions.

## References

- [1] L. Lucchese and S. K. Mitra, "Color image segmentation: a state-of-art survey," *Proceedings of the Indian National Science Academy*, vol. 67-A, pp. 207–221, 2001.
- [2] J. Freixenet, X. Muñoz, D. Raba et al., "Yet another survey on image segmentation: region and boundary information integration," in *Proceedings of the 7th European Conference on Computer Vision (ECCV '02)*, vol. 2352, pp. 408–422, Copenhagen, Denmark, 2002.
- [3] 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.
- [4] 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.
- [5] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888–905, 2000.
- [6] M. Sezgin, O. K. Ersoy, and B. Yazgan, "Segmentation of remote sensing images using multistage unsupervised learning," in *Applications of Digital Image Processing XXVII*, vol. 5558 of *Proceedings of SPIE*, pp. 616–623, Colorado, Colo, USA, August 2004.
- [7] D. M. Ristić, M. Pavlović, and I. Reljin, "Image segmentation method based on self-organizing maps and K-means algorithm," in *Proceedings of the 9th Symposium on Neural Network Applications in Electrical Engineering (NEUREL '08)*, pp. 27–30, Belgrade, Serbia, September 2008.
- [8] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1254–1259, 1998.
- [9] Y. F. Ma and H. J. Zhang, "Contrast-based image attention analysis by using fuzzy growing," in *Proceedings of the 11th ACM International Conference on Multimedia (MM '03)*, pp. 374–381, Berkeley, Calif, USA, November 2003.
- [10] M. L. Gonçalves, M. L. A. Netto, J. A. F. Costa, and J. Zullo Junior, "An unsupervised method of classifying remotely sensed images using Kohonen self-organizing maps and agglomerative hierarchical clustering methods," *International Journal of Remote Sensing*, vol. 29, no. 11, pp. 3171–3207, 2008.
- [11] A. R. F. Araújo and D. C. Costa, "Local adaptive receptive field self-organizing map for image color segmentation," *Image and Vision Computing*, vol. 27, no. 9, pp. 1229–1239, 2009.
- [12] D. E. Ilea and P. F. Whelan, "CTex—an adaptive unsupervised segmentation algorithm based on color-texture coherence," *IEEE Transactions on Image Processing*, vol. 17, no. 10, pp. 1926–1939, 2008.
- [13] V. V. Srinivas, S. Tripathi, A. R. Rao, and R. S. Govindaraju, "Regional flood frequency analysis by combining self-organizing feature map and fuzzy clustering," *Journal of Hydrology*, vol. 348, no. 1-2, pp. 148–166, 2008.
- [14] M. Sujaritha and S. Annadurai, "A three-level clustering algorithm for color texture segmentation," in *Proceedings of the International Conference and Workshop on Emerging Trends in Technology (ICWET '10)*, pp. 645–650, Mumbai, India, 2010.
- [15] G. Wiselin Jiji and L. Ganesan, "A new approach for unsupervised segmentation," *Applied Soft Computing*, vol. 10, no. 3, pp. 689–693, 2010.
- [16] J. Vesanto and E. Alhoniemi, "Clustering of the self-organizing map," *IEEE Transactions on Neural Networks*, vol. 11, no. 3, pp. 586–600, 2000.
- [17] K. Taşdemir and E. Merényi, "Exploiting data topology in visualization and clustering of self-organizing maps," *IEEE Transactions on Neural Networks*, vol. 20, no. 4, pp. 549–562, 2009.
- [18] D. Brugger, M. Bogdan, and W. Rosenstiel, "Automatic cluster detection in Kohonen's SOM," *IEEE Transactions on Neural Networks*, vol. 19, no. 3, pp. 442–459, 2008.



- [19] A. C. Sobieranski, D. D. Abdala, E. Comunello, and A. V. Wangenheim, "Learning a color distance metric for region-based image segmentation," *Pattern Recognition Letters*, vol. 30, no. 16, pp. 1496–1506, 2009.
- [20] K. M. Shaaban and N. M. Omar, "Region-based deformable net for automatic color image segmentation," *Image and Vision Computing*, vol. 27, no. 10, pp. 1504–1514, 2009.
- [21] O. Lézoray and C. Charrier, "Color image segmentation using morphological clustering and fusion with automatic scale selection," *Pattern Recognition Letters*, vol. 30, no. 4, pp. 397–406, 2009.
- [22] H. B. Tan, Z. Q. Hou, X. C. Li, R. Liu, and W. W. Guo, "Improved watershed algorithm for color image segmentation," in *MIPPR 2009: Automatic Target Recognition and Image Analysis*, vol. 7495 of *Proceedings of SPIE*, Yichang, China, October 2009, 74952Z.
- [23] M. Zhang, L. Zhang, and H. D. Cheng, "A neutrosophic approach to image segmentation based on watershed method," *Signal Processing*, vol. 90, no. 5, pp. 1510–1517, 2010.
- [24] T. Celik and T. Tjahjadi, "Unsupervised colour image segmentation using dual-tree complex wavelet transform," *Computer Vision and Image Understanding*, vol. 114, no. 7, pp. 813–826, 2010.
- [25] J. F. Khan, R. R. Adhami, and S. M. A. Bhuiyan, "A customized Gabor filter for unsupervised color image segmentation," *Image and Vision Computing*, vol. 27, no. 4, pp. 489–501, 2009.
- [26] J. H. Park, G. S. Lee, and S. Y. Park, "Color image segmentation using adaptive mean shift and statistical model-based methods," *Computers and Mathematics with Applications*, vol. 57, no. 6, pp. 970–980, 2009.
- [27] 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.
- [28] A. V. Wangenheim, R. F. Bertoldi, D. D. Abdala et al., "Color image segmentation using an enhanced Gradient Network Method," *Pattern Recognition Letters*, vol. 30, no. 15, pp. 1404–1412, 2009.
- [29] L. Garcia Ugarriza, E. Saber, S. R. Vantaram, V. Amuso, M. Shaw, and R. Bhaskar, "Automatic image segmentation by dynamic region growth and multiresolution merging," *IEEE Transactions on Image Processing*, vol. 18, no. 10, pp. 2275–2288, 2009.
- [30] Y. Sun and R. Fisher, "Object-based visual attention for computer vision," *Artificial Intelligence*, vol. 146, no. 1, pp. 77–123, 2003.
- [31] X. Hou and L. Zhang, "Saliency detection: a spectral residual approach," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '07)*, pp. 1–8, June 2007.
- [32] Y. Hu, D. Rajan, and L.-T. Chia, "Robust subspace analysis for detecting visual attention regions in images," in *Proceedings of the 13th Annual ACM International Conference on Multimedia (MULTIMEDIA '05)*, pp. 716–724, Singapore, 2005.
- [33] M. T. López, M. A. Fernández, A. Fernández-Caballero, J. Mira, and A. E. Delgado, "Dynamic visual attention model in image sequences," *Image and Vision Computing*, vol. 25, no. 5, pp. 597–613, 2007.
- [34] T. Avraham and M. Lindenbaum, "Esaliency (extended saliency): meaningful attention using stochastic image modeling," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 4, pp. 693–708, 2010.
- [35] C. Guo and L. Zhang, "A novel multiresolution spatiotemporal saliency detection model and its applications in image and video compression," *IEEE Transactions on Image Processing*, vol. 19, no. 1, pp. 185–198, 2010.
- [36] Y. H. Kuan, C. M. Kuo, and N. C. Yang, "Color-based image salient region segmentation using novel region merging strategy," *IEEE Transactions on Multimedia*, vol. 10, no. 5, pp. 832–845, 2008.
- [37] K. T. Park and Y. S. Moon, "Automatic extraction of salient objects using feature maps," in *Proceedings of the 32nd IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '07)*, vol. 1, pp. 617–620, Honolulu, Hawaii, USA, April 2007.
- [38] D. Mahapatra and Y. Sun, "Using saliency features for graph-cut segmentation of perfusion kidney images," in *Proceedings of the 13th International Conference on Biomedical Engineering (ICBME '08)*, vol. 23, pp. 639–642, Singapore, 2008.
- [39] E. Mendi and M. Milanova, "Image segmentation with active contours based on selective visual attention," in *Proceedings of the 8th WSEAS International Conference on Signal Processing/3rd WSEAS International Symposium on Wavelets Theory and Applications in Applied Mathematics, Signal Processing & Modern Science*, pp. 79–84, Istanbul, Turkey, 2009.
- [40] L. Tan, Y. Cao, M. Yang, and Q. He, "Combined segmentation and visual attention for object categorization and video semantic concepts detection," in *Proceedings of the 3rd International Conference on Pervasive Computing and Applications (ICPCA '08)*, vol. 1-2, pp. 692–697, October 2008.
- [41] J. Zhang, L. Zhuo, and L. Shen, "Regions of interest extraction based on visual attention model and watershed segmentation," in *Proceedings of the IEEE International Conference Neural Networks and Signal Processing (ICNNSP '08)*, vol. 1-2, pp. 375–378, Zhenjiang, China, June 2008.
- [42] K. Fukuda, T. Takiguchi, and Y. Ariki, "Automatic segmentation of object region using graph cuts based on saliency maps and AdaBoost," in *Proceedings of the 13th IEEE International Symposium on Consumer Electronics (ISCE '09)*, vol. 1-2, pp. 36–37, Kyoto, Japan, 2009.
- [43] Y. Fu, J. Cheng, Z. Li, and H. Lu, "Saliency cuts: an automatic approach to object segmentation," in *Proceedings of the 19th International Conference on Pattern Recognition (ICPR '08)*, vol. 1–6, pp. 696–699, Tampa, Fla, USA, December 2008.
- [44] E. Vazquez, T. Gevers, M. Lucassen, J. Van de Weijer, and R. Baldrich, "Saliency of color image derivatives: a comparison between computational models and human perception," *Journal of the Optical Society of America A*, vol. 27, no. 3, pp. 613–621, 2010.
- [45] S. Wan, P. Jin, and L. Yue, "An approach for image retrieval based on visual saliency," in *Proceedings of the International Conference on Image Analysis and Signal Processing (IASP '09)*, pp. 172–175, Taizhou, China, April 2009.
- [46] W. Liu, W. Xu, and H. Li, "A tentative study of visual attention-based salient features for image retrieval," in *Proceedings of the 7th World Congress on Intelligent Control and Automation (WCICA '08)*, vol. 1–23, pp. 7635–7639, Chongqing, China, June 2008.
- [47] A. Perina, M. Cristani, and V. Murino, "Unsupervised learning of saliency concepts for natural image classification and retrieval," in *Proceedings of the 13th Iberoamerican Congress on Pattern Recognition, Image Analysis and Applications*, vol. 5197 of *Lecture Notes in Computer Science*, pp. 169–177, Havana, Cuba, September 2008.

- [48] S. Feng, D. Xu, and X. Yang, "Attention-driven salient edge(s) and region(s) extraction with application to CBIR," *Signal Processing*, vol. 90, no. 1, pp. 1–15, 2010.
- [49] Z. Liu, H. Yan, L. Shen, K. N. Ngan, and Z. Zhang, "Adaptive image retargeting using saliency-based continuous seam carving," *Optical Engineering*, vol. 49, no. 1, Article ID 017002, 2010.
- [50] I. Pirnóg, C. Oprea, and C. Paleologu, "Image content extraction using a bottom-up visual attention model," in *Proceedings of the 3rd International Conference on Digital Society (ICDS '09)*, pp. 300–303, Cancun, Mexico, February 2009.
- [51] S. Wan, Y. Xiong, and L. Yue, "An algorithm of wavelet image fusion based on visual saliency," *Chinese Journal of Electronics*, vol. 17, no. 2, pp. 237–241, 2008.
- [52] R. Hong, C. Wang, M. Wang, and F. Sun, "Saliency preserving multifocus image fusion with dynamic range compression," *International Journal of Innovative Computing, Information and Control*, vol. 5, no. 8, pp. 2369–2380, 2009.
- [53] Q. Ma and L. Zhang, "Image quality assessment with visual attention," in *Proceedings of the 19th International Conference on Pattern Recognition (ICPR '08)*, vol. 1–6, pp. 2783–2786, Tampa, Fla, USA, December 2008.
- [54] T. Kohonen, *Self-Organizing Maps*, Springer, 3rd edition, 2001.
- [55] J. B. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, pp. 281–297, 1967.
- [56] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 1, no. 2, pp. 224–227, 1979.
- [57] M. Halkidi and M. Vazirgiannis, "Clustering validity assessment using multi representatives," in *Proceedings of the Hellenic Conference on Artificial Intelligence (SETN '02)*, Thessaloniki, Greece, 2002.
- [58] A. R. S. Marçal and J. S. Borges, "Estimation of the "natural" number of classes of a multispectral image," in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS '05)*, vol. 6, pp. 3788–3791, Seoul, Korea, July 2005.
- [59] N. R. Pal and J. Biswas, "Cluster validation using graph theoretic concepts," *Pattern Recognition*, vol. 30, no. 6, pp. 847–857, 1997.
- [60] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, John Wiley & Sons, 2nd edition, 2001.
- [61] H. Zhang, J. E. Fritts, and S. A. Goldman, "An entropy-based objective evaluation method for image segmentation," in *Storage and Retrieval Methods and Applications for Multimedia 2004*, vol. 5307 of *Proceedings of SPIE*, pp. 38–49, January 2004.
- [62] M. Borsotti, P. Campadelli, and R. Schettini, "Quantitative evaluation of color image segmentation results," *Pattern Recognition Letters*, vol. 19, no. 8, pp. 741–747, 1998.
- [63] R. C. Gonzales and R. E. Woods, *Digital Image Processing*, Pearson Education, 3rd edition, 2008.
- [64] <http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/>.
- [65] <http://www.cis.hut.fi/projects/somtoolbox/>.
- [66] A. Hanbury and B. Marcotegui, "Morphological segmentation on learned boundaries," *Image and Vision Computing*, vol. 27, no. 4, pp. 480–488, 2009.



**Hindawi**

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

