Iterative Optimization Scheme for Image Segmentation

M. S. Karande
Lecturer-IT Dept
K. K. Wagh Polytechnic, Nashik

D. B. Kshirsagar
HOD -Computer Engg. Dept.
SRES COE, Kopargaon

ABSTRACT
Image Segmentation is an integral part of computer vision. In this paper image segmentation is formulated as label relabeling problem under probability framework. To estimate the label configuration, an iterative optimization scheme is proposed to alternately carry out the maximum a posteriori (MAP) estimation and the maximum likelihood (ML) estimation. This algorithm can automatically partition the image into regions without human intervention. The segmentation obtained is very close to human perception. Comparing to other state-of-the-art algorithms, extensive experiments have shown that this algorithm performs the best.

General Terms
Pattern Recognition.

Keywords
Image Segmentation, Maximum a Posteriori, Maximum Likelihood

1. INTRODUCTION
Segmentation is the process of reducing an image to regions that correspond to structural units, or to some specific property. Segmentation may be manual, semi-automatic or automatic depending on the complexity of the task. The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics. The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image. Each of the pixels in a region is similar with respect to some characteristic or computed property, such as color, intensity or texture. Adjacent regions are significantly different with respect to the same characteristics.

Although the problem of image segmentation has been studied for more than three decades, no single segmentation technique has provided satisfactory result for all applications.

2. Related Work
Available image segmentation algorithms can be classified into two groups: contour-based approaches and region-based approaches. Contour-based approaches try to find the boundaries of objects in an image, while region-based approaches attempt to split an image into connected regions. The main idea of contour-based approaches is to start with some initial boundary shape represented in the form of a spline curve, and iteratively modifies it by shrink and expansion operations to minimize some energy function. One problem existing in these algorithms is that they are easy to get trapped in local minima. In addition, they need manually specified initial curves close to the objects of interest.

Region-based approaches try to classify an image into multiple consistent regions or classes. Thresholding is the simplest segmentation method but its performance is usually far from satisfactory. Watershed segmentation [8] is one of the traditional region-based approaches. The watershed transform is often used to segment touching objects. It finds intensity valleys in an image if the image is viewed as a surface with mountains (high intensity regions) and valleys (low intensity regions). Morphological operations are always used to handle the over-segmented problem in the output obtained by the watershed transform. Usually, watershed is used for the segmentation of foreground and background (two-class) of an image. For a general color image with many different regions, it often gives a bad result.

The K-means algorithm [1] is the basic one. However, the K-means is not good enough because it does not take account of the spatial proximity of pixels. It is, thus, often used in the initialization step for other approaches. Expectation-maximization (EM) [2] performs segmentation by finding a Gaussian mixture model in an image feature space. One short coming of EM is that the number of regions is kept unchanged during the segmentation, which often causes wrong results because different images usually have different numbers of regions. Theoretically, the minimum description length (MDL) principle [2] can be used to alleviate this problem, but the segmentation has to be carried out many times with different region numbers to find the best result. This takes a large amount of computation, and the theoretically best result may not accord with this perception. In [3], a mean shift algorithm is proposed for image segmentation. Mean shift is a nonparametric clustering technique which neither requires to know the number of clusters in advance nor constrains the shapes of the clusters However, it often obtains over-segmented results for many natural images.

Recently, a number of graph-based approaches are developed for image segmentation. Shi and Malik's [4] normalized cuts are able to capture intuitively salient parts in an image. Normalized cuts are a land mark in current popular spectral clustering research, but it is not perfectly fit to the nature of image segmentation because adhoc approximations must be introduced to relax the NP-hard computational problem. These approximations are not well understood and often lead to unsatisfactory results. In addition, the heave computational cost is a disadvantage in spectral clustering algorithms.

Another popular segmentation approach based upon MRFs is graphcut algorithm [5]. This algorithm relies on human interaction, and solves the two-class segmentation problem only, i.e., separating an image into only background and object regions, with some manually given seed points.

3. Outline of the Work

This paper discusses a new image segmentation algorithm based upon a probability maximization model. An iterative optimization scheme consisting of the maximum a posteriori (MAP) and the maximum likelihood estimation is carried out alternately. This paper discusses a novel probabilistic model and uses graph cuts to solve the multiple region segmentation problems with the number of regions automatically adjusted according to the properties of the regions. This algorithm can cluster relevant regions in an image well, with the segmentation boundaries matching the region edges. Extensive experiments show that this algorithm can obtain results highly consistent with human perception. The rest of this paper is organized as follows. In Section 4, probabilistic model is discussed and Section 5 discusses the detail of algorithm. Section 6 discusses the performance of this algorithm. Section 7 concludes the paper.

4. Probabilistic Model

For a given image, the features of every pixel are expressed by a 4-D vector

$$I(p)=(I_L(p), I_a(p), I_b(p), I_t(p))^T \qquad (1)$$

where $I_L(p)$, $I_a(p)$, $I_b(p)$ are the components of p in the L*a*b*color space, and $I_t(p)$ denotes the texture feature of p. In this paper, the texture contrast defined in [2] (scaled from [0, 1] to [0,255]) is chosen as the texture descriptor. Figure 1 shows an example of the features.

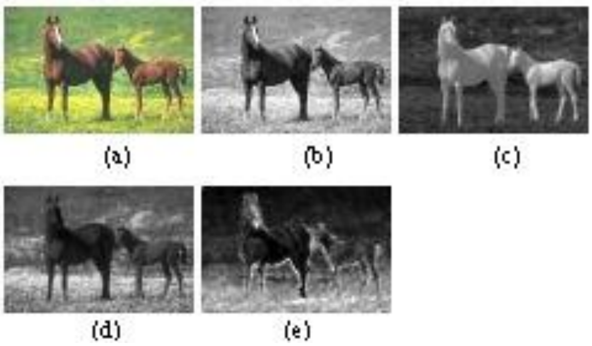

Figure 1(a) Original color image.(b)–(d)Three components of (a) in L*a*b* colorspace. (e) Texture contrast of (a).

The task of image segmentation is to group the pixels of an image into relevant regions. If the problem is formulated it as a labeling problem, the objective is then to find a label configuration $f = \{f_p|p\}$ where f_p is the label of pixel denoting which region this pixel is grouped into. Generally speaking, a “good” segmentation means that the pixels within a region i should share homogeneous features represented by a vector $\varphi(i)$ that does not change rapidly except on the region boundaries. The introduction of $\varphi(i)$ allows the description of a region, with which high level knowledge or learned information can be incorporated into the segmentation. Suppose that there are k possible region labels. A 4-D vector

$$\varphi(i) = (\bar{I}_L(i), \bar{I}_a(i), \bar{I}_b(i), \bar{I}_t(i))^T \qquad (2)$$

is used to describe the properties of label(region), where the four components of $\varphi(i)$ have the similar meanings to those of the corresponding four components of I(p). Let $\varphi = \{\varphi(i)\}$ be the union of the region features. If P and Φ are known, the segmentation is to find an optimal label configuration $\hat{f}$, which maximizes the posterior possibility of the label configuration

$$\hat{f} = \arg\, max \;\; \Pr(f\,|\,\varphi, P) \qquad (3)$$

where Φ can be obtained by either a learning process or an initialized estimation. However, due to the existence of noise and diverse objects in different images, it is difficult to obtain Φ that is precise enough. Thus, an iterative method is used to solve the segmentation problem.

Suppose that φ^n and f^n are the estimation results in the nth iteration. Then the iterative formulas for optimization are defined as

$$f^{n+1} = \arg\; \max_{f} \Pr(f|\varphi^n, P) \qquad (4)$$

$$\varphi^{n+1} = \underset{f}{\arg\max} \Pr(f^{n+1}|\varphi, P) \qquad (5)$$

This iterative optimization is preferred because (4) can be solved by the MAP estimation, and (5) by the ML estimation.

5. Segmentation Algorithm

The MAP estimation is used to detect the edges of the image and the color space is used to segment the images by colors. The K-means algorithm is used for initializations of the regions. The texture contrast is used to segment the outline of the each object in the image and labeling is used to delete the unwanted portion of the image and segment the each object by each color. The MAP-ML is used to segment the image by each object in the same image. The graph cut algorithm is an unsupervised algorithm used for over segmentation and computation problem.

Algorithm: MAP-ML Image Segmentation: **[6]**
Input: an RGB color image.

Step 1: Convert the image into L*a*b* space and calculate the texture contrast [2].

Step 2: Use the K-means algorithm to initialize features of region Φ.

Step 3: Iterative optimization.

3.1: MAP estimation— Estimate the label configuration f based upon current Φ using the graph cut algorithm [7].

3.2: Relabeling—Set a unique label to each connecting region to form a new cluster, obtaining a new f

3.3: ML estimation—Refine Φ based upon current f

Step 4: If Φ and f do not change between two successive iterations or the maximum number of iterations is reached, go to the output step; otherwise, go to step 3.

Output: Multiple segmented regions of the image.

After step 3.1, it is possible that two non adjacent regions are given the same label. For example, the upper-left and the lower-right regions in Figure 2(a) are both labeled by 1. After step 3.2, each of the connected regions has a unique label [see Figure 2(b)]. The MAP estimation is an NP-hard problem. Boykov et al. [7] proposed to obtain an approximate solution via finding the minimum cuts in a graph model. Minimum cuts can be obtained by computing the maximum flow between the terminals of the graph. In [7], an efficient max-flow algorithm is given for solving the binary labeling problem. In addition, an algorithm, called α expansion with the max-flow algorithm embedded, is presented to carry out multiple labeling iteratively. In this algorithm, the α expansion algorithm is used to perform step 3.1.

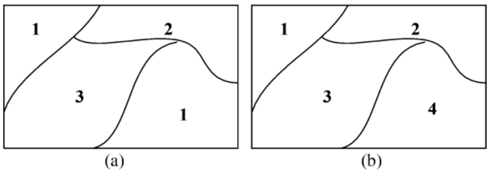

Figure 2 Relabeling of the regions. (a) Result before the relabeling (b) Result after the relabeling

One remarkable property of this algorithm is the ability to adjust the region number automatically during the iterative optimization with the relabeling step embedded in to the MAP and ML estimations. Figure 3 gives an example to show how the iterations improve the segmentation results. Comparing Figure 3(b)–(d), it is seen that the final result is the best.

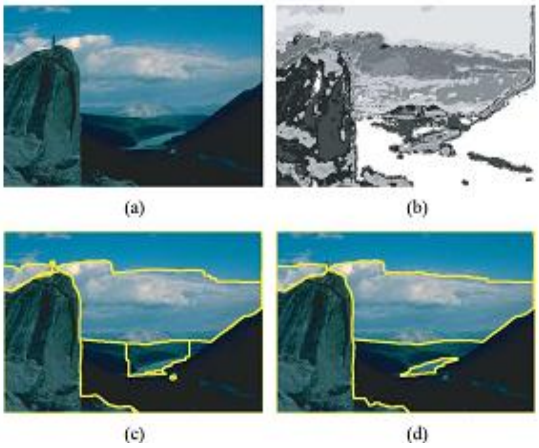

Figure 3 Segmentation example. (a) Original image. (b) Result of initial K-means clustering with K=10 (c) Result of the first iteration with K adjusted to 8 automatically. (d) Converged result after 4 iterations with K changed to 6 automatically

6. Experimental Results

The MAP-ML algorithm is tested on the Berkeley benchmark for evaluating segmentation algorithms and compares the results with those obtained by state-of-the-art image segmentation algorithms. The Berkeley database contains 300 natural images of size 321 x 481 (or 481 x 321), with ground truth segmentation results obtained from human subjects.

The compared algorithms in these experiments include: normalized cuts (NC) [4] and Mean Shift (MS) [3]. In this algorithm, the initial cluster number in the K-means algorithm is set to10 and the smoothness factor c to 4000. The region number in NC is set to 20, which is the average number of segments marked by the human subjects in each image. Since NC cannot handle an image of size 321x481(or 481 x 321) due to the overflow of the memory, all the input images for them are shrunk into a size 150 x 150, and the segmentation results are enlarged to their original sizes.

6.1 Qualitative Comparisons

The segmentation results obtained by the three algorithms are shown in Figure 4. From these examples, the following observations are seen. NC tends to partition an image into regions of similar sizes, resulting in the region boundaries different from the real edges. MS give strongly over-segmented results. Compared with these other algorithms, it is easy to see that MAP-ML algorithm obtains the best results, in which the generated boundaries match the real edges well and the segmented regions are in accordance with human perception. This algorithm can adapt the number of regions to different images automatically although all the initial numbers in the initialization step are set to 10.

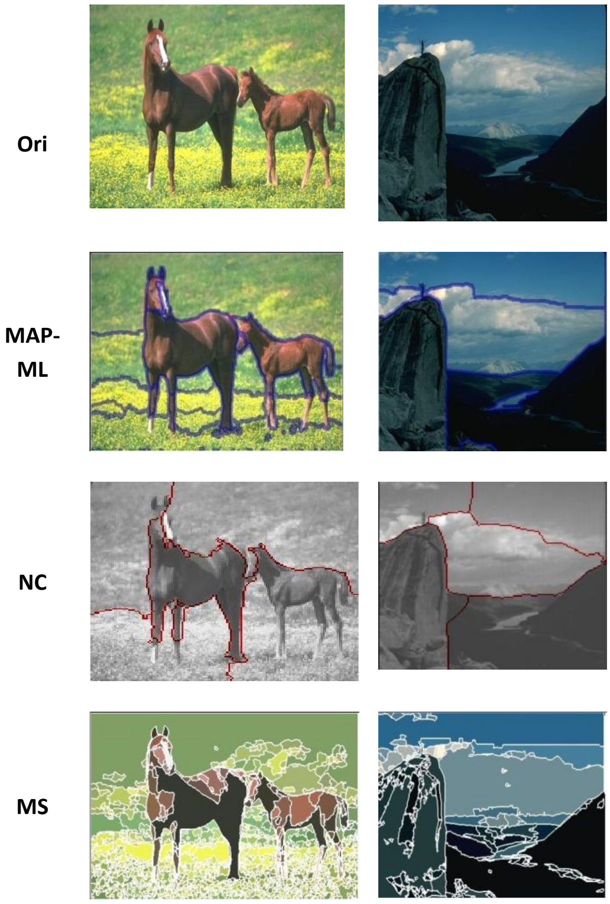

Figure 4 Segmentation results on the images.

It is clearly observed that MAP-ML algorithm performs the best.

7. CONCLUSION

In this paper, a novel image segmentation algorithm is discussed. The algorithm is formulated as a labeling problem using a probability maximization model. An iterative optimization technique combining the MAP and ML estimations is employed in this framework. Under the Gaussian model, the MAP estimation problem is solved using graphcut and the ML estimation is obtained by finding the means of the region features. This algorithm is compared with other state-of-the- art image segmentation algorithms. The qualitative results demonstrate that this algorithm outperforms the others.

8. REFERENCES

[1] R. Duda, P. Hart, and D. Stork, Pattern Classification, 2nd ed. Hoboken, NJ: Wiley, 2001.

[2] C. Carson, S. Belongie, H. Greenspan, and J. Malik, "Blobworld: Image segmentation using expectation-maximization and its application to image querying," IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 8, pp. 1026–1038, Aug. 2002.

[3] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 5, pp. 603–619, May 2002.

[4] J. Shi and J. Malik, "Normalized cuts and image segmentation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 8, pp. 888–905, Aug. 2000.

[5] V. Kolmogorov and R. Zabih, "What energy functions can be minimized via graph cuts?," IEEE Trans. Pattern Anal. Mach. Intell., vol. 26, no. 2, pp. 147–159, Feb. 2004.

[6] Shifeng Chen, Liangliang Cao, Yueming Wang, Jianzhuang Liu, "Image Segmentation by MAP-ML Estimations," IEEE Trans. on Image Processing, Vol. 19, No. 9, pp. 2254-2264, September 2010.

[7] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," IEEE Trans. Pattern Anal. Mach. Intell., vol.23, no. 11, pp. 1222–1239, Nov.2001.

[8] L. Vincent and P. Soille, "Watersheds in digital spaces: An efficient algorithm based on immersion simulations," IEEE Trans. Pattern Anal. Mach. Intell., vol.13, no. 6, pp. 583-598, Jun.1991.