TOP-DOWN IMAGE SEGMENTATION USING THE MUMFORD-SHAH FUNCTIONAL AND LEVEL SET IMAGE REPRESENTATION

Yongsheng Pan

Scientific Imaging and Computing Institute
University of Utah

ABSTRACT

A top-down image segmentation method is proposed in this paper, utilizing level set image representation and the piecewise-constant Mumford-Shah functional. The method achieves top-down hierarchical segmentation by taking advantage of the tree structure provided by level set image representation. The piecewise-constant Mumford-Shah functional is utilized in the proposed method to determine if each node in the tree segments the image. Experimental results show that this method is able to segment complicated real images.

***Index Terms*—** Image segmentation, the Mumford-Shah functional, level set image representation, tree of shapes, region merging.

1. INTRODUCTION

Image segmentation is one of the most important topics in image processing. It is designed to divide an image into several regions based on image content. The Mumford-Shah functional [1] is one of the most popular models for image segmentation. The minimization of this functional tends to find a segmented image which is close to the original image while containing low variation and short object boundaries.

Curve evolution methods [2] [3] are usually utilized for the minimization of the Mumford-Shah functional. These methods have solid theoretical foundations and stable implementations. However, they are usually computationally intense. Furthermore, they can only work at cartoon-like simple images. Their applications seem to be limited for complicated real images.

A more practical image segmentation method is proposed in [4] based on bottom-up region growing to minimize the piecewise-constant Mumford-Shah functional. This method takes every pixel in an image as a region, and it merges two regions if this merging decreases the discrete-version of the piecewise-constant Mumford-Shah functional. This method is shown to be efficient for complicated images.

Top-down image segmentation methods, however, have advantages in many cases. They are less likely to be trapped in local minimum, and they are more robust to noise. Therefore, it may be beneficial to extend the method in [4] in a top-down hierarchical framework.

A top-down image segmentation method is proposed in this paper, in which an image is represented using level sets. A level set corresponds to the region of a connected component (e.g., $\{$(x,y) : I(x,y)$\leq$50 $\}$) in an image I. A shape is further defined to be a connected component with its holes filled. This representation describes an image as a tree of shapes instead a set of pixels. The tree structure is utilized in the proposed method for top-down image segmentation.

The proposed method assumes that the segmentation results of an image correspond to boundaries of the shapes in the tree structure. Each shape in the tree is examined from top level to lower level to see if it should be segmented based on the piecewise-constant Mumford-Shah functional. A shape is segmented if its segmentation decreases the functional. This top-down segmentation method is able to segment complicated real images.

This paper is organized as follows. Section 2 provides background information about the Mumford-Shah functional and level set image representation. Details of the proposed method are introduced in Section 3. Experimental results are shown in Section 4. Section 5 contains conclusion and future work.

2. BACKGROUND

The Mumford-Shah functional [1] is first introduced in this section. Background information on level set image representation is then introduced.

2.1. The Mumford-Shah Functional

Let I_0 be a function representing the image to be segmented and I be a function representing the segmented image. Both I_0 and I are defined on a planar domain R. Let R_i be disjoint connected open subsets of R with piecewise-smooth boundaries, and let Γ be the union of the portions of the boundaries of R_i inside R. Then the Mumford-Shah functional is defined

978-1-4244-2354-5/09/$25.00 ©2009 IEEE

as

$$E(I,\Gamma) = \mu \iint_R (I - I_0)^2 dxdy + \iint_{R-\Gamma} \|\nabla I\|^2 dxdy + \nu|\Gamma| \quad (1)$$

where $|\Gamma|$ represents the total length of Γ, and μ and ν are positive constants. The first term in (1) penalizes the integrated squared error between I and I_0; the second term penalizes variation of I within each region; the third term penalizes the boundary length. All three terms work together to make the functional meaningful.

The functional in (1) admits piecewise-smooth solutions. In most cases, a special case of (1), in which I is restricted to be piecewise-constant, is applied. This special case of the Mumford-Shah functional takes the following form

$$E(\Gamma) = \Sigma_i \iint_{R_i} (I_0 - mean_{R_i}(I_0))^2 dxdy + \nu|\Gamma| \quad (2)$$

The functional in (1) is usually hard to be minimized in real applications. So the piecewise-constant functional in (2) is utilized in this paper.

2.2. Level Set Image Segmentation

Image representation using a tree of shapes [5] [6] utilizes the inferior or the superior of a level line to represent an object. This representation also provides a tree structure to represent the spatial relationship for the objects in an image.

For a gray image $I : \Omega \to \mathbb{R}$ with $\Omega \subset R^2$, the upper level set χ_λ of value λ and the lower level set χ^μ of value μ are defined in [5] as $\chi_\lambda = \{x \in R^2, I(x) \geq \lambda\}$ and $\chi^\mu = \{x \in R^2, I(x) \leq \mu\}$.

The above representations are complete for images, which means that the family of the upper level sets χ_λ (or the family of the lower level sets χ^μ) is sufficient to reconstruct the image [6] because of the following relationship [5]: $I(x) = \sup\{\lambda | \mathrm{x} \in \chi_\lambda\} = \inf\{\mu | \mathrm{x} \in \chi^\mu\}$.

Note that the geometrical inclusion holds for the level sets. The family of upper (lower) level sets is decreasing (increasing) because [5] $\lambda \leq \mu \Rightarrow \chi_\lambda \supset \chi_\mu \quad \text{and} \quad \chi^\lambda \subset \chi^\mu$

The nesting of level sets provides an inclusion tree for an image. Each node in the tree is called a shape, which is defined as the connected components of a level set and the holes inside them. Fig. 1 shows an example of the tree of shapes generated for a synthetic image. A tree of shapes shown in Fig. 1(h) is constructed for the image in Fig. 1(a). The whole image acts as the root of the tree, which locates at the top level. The shapes in the same level are spatially disjoint in the image. The shapes in the lower level are spatially included in the shapes in the next higher level. The tree of shapes, therefore, provides a natural way to represent the spatial relationships between the shapes in the image.

It is straight forward to find upper level sets and lower level sets in a gray image by thresholding. The total order (or lexicographical order) proposed in [7] extends level set representation to color images. A tree of shapes can be further constructed by the nesting of level sets. However, this method is computationally intense. The fast level lines transform (FLLT) [6] provides a faster way to construct a tree of shapes.

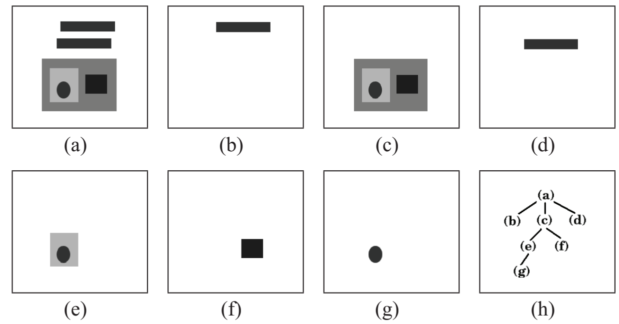

Fig. 1. Illustration of the tree of shapes. (a) The original image, which locates at the root of the inclusion tree. (b)(c)(d) Shapes in the first layer of the tree of shapes. (e)(f) Shapes in the second layer of the tree of shapes. (g) Shapes in the third layer of the tree of shapes. (h) The Structure of the tree of shapes.

3. PROPOSED METHOD

Details of the proposed method are introduced in this section. The proposed method assumes that the segmentation results correspond to boundaries of the shapes in the tree structure. It utilizes the tree of shapes for top-down segmentation, and it utilizes the Mumford-Shah functional in (2) for the segmentation of each shape in the tree structure.

The tree structure introduced by level set image representation in Section 2.2 provides a top-down hierarchy for complicated image segmentation. In the proposed method the shapes in the tree are processed from top level to low level. Suppose the shape in Fig. 1(g) is chosen to be processed first. All its parent shapes in Fig. 1(e), Fig. 1(c) and Fig. 1(a) are checked if they have been processed. These parent nodes are processed from top level to lower level before the shape in Fig. 1(g) can be processed. This process is illustrated in Fig. 2(b) - Fig. 2(d) which correspond to the first three processed shapes of the proposed method. This top-down hierarchy makes sure that each shape in the tree can be processed only if its parent shape has been processed.

The piecewise-constant Mumford-Shah functional in (2) is utilized to determine if a shape should be segmented or not. If the segmentation of a shape decreases the functional, then it will be segmented. Otherwise it will not be segmented, and it will be merged to its parent shape. The procedure continues until all the shapes in the tree are processed.

The piecewise-constant Mumford-Shah functional in (2) segments an image into piecewise-constant regions [1].

Therefore, the mean value and the area of the region contained in each shape need to be calculated first in the implementation. A recursive function is designed here to go through all the shapes in a bottom-up way to calculate their mean values. When a shape is segmented, the mean value and the area of its parent shape will be updated. When a shape merges with its parent shape, this child shape will be removed from the tree. The mean value and the area of its parent shape will be updated to be the average of both shapes.

4. EXPERIMENTAL RESULTS

Experimental results are shown in this section. The construction of the tree of shapes for an image (FLLT) is accomplished using the megawave software [8]. All the experiments are performed on a Dell XPS 720 with Intel Core2 processor Q6600 (2.4GHZ) and 4GB DDR2 memory. This section only show experimental results from the first 10 segmentations for complicated real images.

One practical issue is that an image may contain hundreds of thousands of shapes. For instance the image in Fig. 4(a) contains 27304 shapes in its tree structure. It can be expected that the child shapes of a segmented shape will also be segmented if there are no big differences between them. These child shapes may be redundant for real applications. So the shapes whose area are very similar to their parent shapes are not considered for segmentation. The similarity here depends on applications. In the implementation a shape is considered for segmentation only if its area lies between 10%-90% of its parent shape. This provides some flexibility in real applications. For example, the fourth segment shown in Fig. 6(d) will be segmented as the second if the range is shrinked from 10%-90% to 30%-70%. In this case the shape in Fig. 6(c) in not considered for segmentation.

Fig. 2 demonstrates the proposed method using a synthetic image in Fig. 1. It can be seen that the shapes are segmented in a top-down hierarchy following to the tree structure in Fig. 1. It means that a shape in the tree can be processed only after its parent shape has been processed, which is shown in Fig. 2(b) 2(c) 2(d). All the objects have been segmented after 6 segmentations.

Fig. 3 shows segmentation results for another synthetic image. Fig. 3(b)-Fig. 3(g) show the segmentation results from the 1st, 2nd, 3rd, 4th, 6th, and 8th iteration respectively. The shapes are segmented in a top-down hierarchy as in Fig. 2. All regions have been segmented as shown in Fig. 3(h).

Fig. 4 shows experimental results for a real image. Fig. 4(b)-Fig. 4(e) shows the first several segments, which are meaningful regions from the image. Fig. 4(g) shows the segmentation result from the Chan-Vese model in [3], and Fig. 4(h) shows the results from the bottom-up method in [4]. It can be seen that the proposed method is more efficient than the Chan-Vese model. It provides more meaningful segments than the bottom-up method in [4] although it is less efficient.

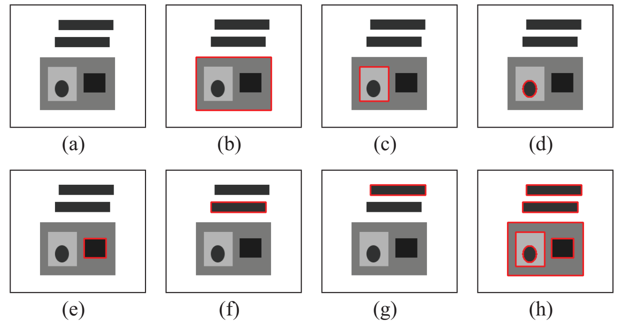

Fig. 2. Demonstration of the proposed method using a synthetic image. (a) The original synthetic image (200 * 200). (b)-(g) The first 6 segments from (a). (h) Final segmentation results. $\nu = 0.01 * 255^2$. Time: 0.093 sec.

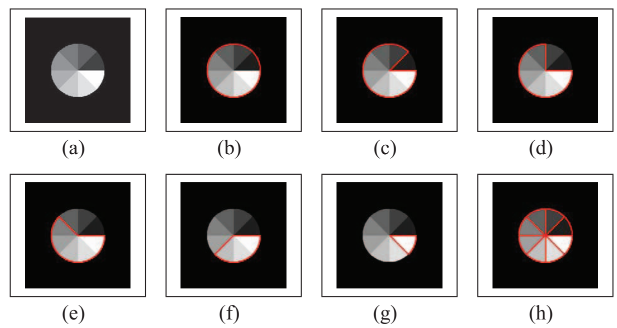

Fig. 3. Demonstration of the proposed method using another synthetic image. (a) The original image (256 * 256). (b)-(g) The 1st, 2nd, 3rd, 4th, 6th, and 8th segments from (a). (h) Final segmentation results. $\nu = 0.01 * 255^2$. Time: 0.144s.

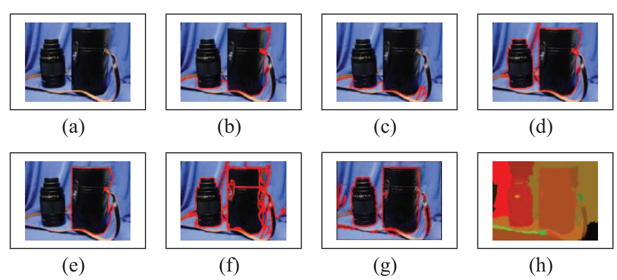

Fig. 4. Real image segmentation using the proposed method. (a) The original image (300 * 225). (b)-(e) The 1st, 3rd, 7th and 8th segmentation. (f) Final segmentation results from the first 10 segments. $\nu = 100 * 255^2$. Time: 5.37s. (g) Segmentation results using Chan-Vese curve evolution [3]. Time: 25.87s. (h) Segmentation results using the bottom-up segmentation method in [4]. Time: 2.66s.

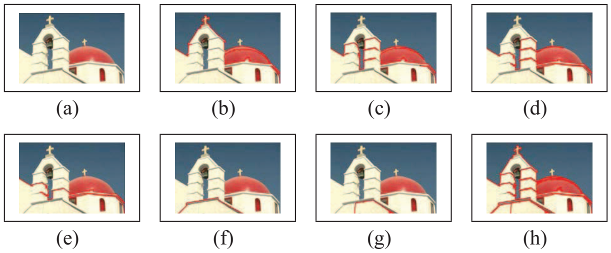

Fig. 5. Real image segmentation using the proposed method. (a) The original image (481 * 321). (b)-(g) The 1st, 2nd, 3rd, 4th, 5th, and 7th segments from (a). (h) Final segmentation results from the first 10 segments. $\nu = 100 * 255^2$. Time: 9.81s.

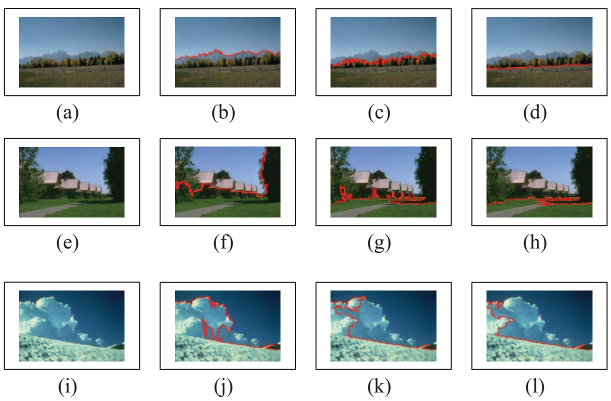

Fig. 6. Real image segmentation using the proposed method. (a) An original image (481 * 321). (b)-(d) The 1st, 3rd and 4th segments from (a). $\nu = 0.1 * 255^2$. Time: 13.10s.(e) An original image (481 * 321). (f)-(h) The first 3 segments from (e). $\nu = 100 * 255^2$. Time: 3.37s. (i) An original image (481 * 321). (j)-(l) The 1st, 4th and 5th segments from (i). $\nu = 100 * 255^2$. Time: 12.73s.

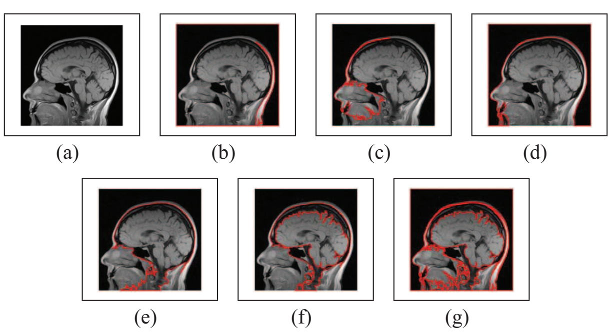

Fig. 7. Real image segmentation using the proposed method. (a) An original medical image (512 * 512). (b)-(f) The first 5 segments from (a). (g) Final segmentation results. $\nu = 255^2$. Time: 2.62s.

Fig. 5-Fig. 7 show more experimental results for real images. Several observations can be acquired from these results. First, the proposed method gives reasonable segmentation for real images. Second, the proposed method is efficient in the sense that it provides multiple segments in a relatively short time. Third, post-processing may be necessary to select the best segments. The utilization of prior knowledge may also improve the proposed method. Fourth, the acquired object boundaries are not regularized very well. This comes from the formulation of the proposed method, and it will be examined in later publications.

5. CONCLUSION

A top-down image segmentation method is proposed in this paper. It utilizes level set image representation and the piecewise-constant Mumford-Shah functional. It is shown to be able to segmented complicated real images. Future research will be focused on the regularization of the object boundaries.

6. REFERENCES

[1] D. Mumford and J. Shah, "Optimal approximation by piecewise smooth functions and associated variational problems," *Comm. On Pure and Appl. Math.*, vol. XLII, pp. 577–685, 1989.

[2] J. Sethian, *Level Set Methods and Fast Marching Methods*, Cambridge University Press, 1999.

[3] T.F. Chan and L.A. Vese, "Active contours without edges," *IEEE Trans. Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.

[4] Y. Pan, J.D. Birdwell, and S.M. Djouadi, "Bottom-up hierarchical image segmentation using region competition and the Mumford-Shah functional," in *Proceedings of International Conference on Pattern Recognition*, Hong Kong, China, Aug. 2006, pp. 117–121.

[5] V. Caselles, "Topographic maps and local contrast changes in natural images," *International Journal of Computer Vision*, vol. 33, no. 1, pp. 5–27, 1999.

[6] P. Monasse and F. Guichard, "Fast computation of a contrast-invariant image representation," *IEEE Transactions on Image Processing*, vol. 9, no. 5, pp. 860–872, 2000.

[7] B. Coll and J. Froment, "Topographic maps of color images," in *Proceedings of International Conference on Pattern Recognition*, Barcelona, Spain, Sep. 2000, vol. 3, pp. 613–616.

[8] J. Froment, L. Moisan, and J-M. Morel, "the megawave package," *http://megawave.cmla.ens-cachan.fr*.