Segmenting an object on a uniform background: comparison of segmentation algorithms

Ashok Yogi[1], Matjaž Majnik[2], Danijel Skočaj[2]

[1]*Department of Electronics and Electrical Engineering, IIT-Guwahati, Guwahati(Assam) 781 039, India*
[2]*University of Ljubljana, Faculty of Computer and Information Science, Večna pot 113, 1000 Ljubljana, Slovenia*
E-mail: s.yogi@iitg.ernet.in, matjaz.majnik@fri.uni-lj.si, danijel.skocaj@fri.uni-lj.si

Abstract

Image segmentation plays a crucial role in many tasks requiring image analysis. In this paper we study six image segmentation algorithms and analyze their performance in a specific task of segmenting a single object on a uniform background in different settings and illumination conditions.

1 Introduction

Image segmentation is one of the most important tasks in computer vision. The main goal of image segmentation is to group together image pixels that belong to the same entity in an image, therefore to segment the image to corresponding regions.

In general, segmentation is an ill-defined problem. General images can be segmented in different ways, up to a different level of detail, depending on the task. In this paper we will limit our research to a specific, more constrained problem: segmentation of a single object placed on a uniform background. We will therefore deal with the scenarios, where

- the background is of uniform color, without the texture and clutter,
- there is only one object on a scene, therefore no occlusions (except self-occlusions) may occur,
- the object is positioned in the center of the image and is fully contained in the image,
- the object can be of arbitrary colors and shape,
- the illumination conditions are not constrained.

The goal is well-defined: to partition an image into two regions containing pixels that belong to the object and to the background, respectively.

Such constrained settings often simplify the segmentation problem. Such an example is depicted in Fig. 1(a), where the background is very homogeneous and the boundary between the object and the background is clear and sharp. However, when the same object in the same position is illuminated with a strong point light source, such as in the image depicted in Fig. 1(b), the segmentation task becomes significantly more difficult due to nonhomogeneous background and strong shadows. Even more difficult problem we face when the object to be segmented, or a part of it, is of a similar color as the background (e.g., Fig. 1(c)).

(a) (b) (c)

Figure 1: Examples of objects with (a) uniform illumination, (b) strong shadows, (c) problematic object color.

In machine vision applications, the problems depicted in Figs. 1(b) and (c) are usually avoided by assuring suitable illumination conditions and carefully choosing the color of the background. We will refer to scenarios, where this is not possible (e.g., where the illumination cannot be controlled). The goal of this paper is to analyze the performance of several segmentation algorithms in such scenarios. We will briefly describe six chosen segmentation algorithms and analyze their performance in different conditions and illumination settings.

The paper is organized as follows. In Section 2 we describe the algorithms that were used in the comparison. In Section 3 we present the experimental results. Finally, we conclude the paper in Section 4.

2 Methods

Contour evolution-based image segmentation methods are well-suited for segmenting non-textured images with only one object. This paper attempts to compare the performance of Active contour [1], Chan Vese model [2], DRLSE [3] and GrabCut [4] under different illumination conditions. Otsu's segmentation method [5] and adaptive thresholding [6] are included for comparison reference.

2.1 Active Contour Model

Active contour model (also known as *snakes*) [1] has been proven to be a promising framework for image segmentation. The fundamental idea of the active contour is to evolve a spline curve under the influence of energy functions. Energy functions are set up in such a way that in the equilibrium position the spline curve conforms to the object boundary or to other desired features. In steady state, the internal energy term (controlling the smoothness of the spline curve) counterbalances the external energy term (edge image), thereby conforming the curve at the edges. However, the sensitivity of snakes to initialization and poor convergence have limited its use in image segmentation. To overcome these limitations, various improvements to the active contour method have been proposed in the past [7, 8].

2.2 Chan Vese Model

Level set method, based on the active contour model, involves representing contours as the zero level of an implicit, level set function [9]. The level set method overcomes the limitation of contour continuity in the active snakes model, thereby handling topological changes like splitting and merging. Chan and Vese proposed a model of active contour without edges (Chan Vese model) [2], which is based on level set evolution and Mumford-Shah segmentation techniques [10]. Unlike in active contours, the curve evolution stopping term in the Chan Vese model is based on the minimization of Mumford-Shah functional [10]. Consequently, the Chan Vese model is capable of segmenting objects with minimal or no gradient at all. However, as the Chan Vese model is based on global image information, it fails to segment images with intensity inhomogenity. Another drawback of Chan Vese model is its slow convergence due to the step of reinitialization. Reinitialization involves assigning a signed distance function to the downgraded level set, thereby preventing any irregularities in the curve evolution.

2.3 DRLSE

Li and Xu redress the constraint of reinitialization by proposing distance regularized method for level set evolution (DRLSE) [3]. In DRLSE, the energy functional for curve evolution inherently contains the distance regularization term along with the external energy, thereby eliminating the need for reinitialization. As the distant regularization term is embedded in the level set evolution, the DRLSE method does not require reinitialization at periodic intervals, thereby making the method computationally fast and accurate. However, unlike the Chan Vese method, DRLSE fails to detect objects with weak boundaries, resulting in boundary leakage.

2.4 GrabCut

GrabCut is an interactive object segmentation method based on the application of graph cuts in conjunction with some apriori knowledge of foreground and background of an image [4]. GrabCut uses the graph cut algorithm to solve an optimization problem (Energy cost function) [11] by creating a specific weighted graph model, where each vertex corresponds to an image pixel and weights of each edge, connecting vertices, represent similarity between pixels. The main problem of GrabCut is its inability to segment images with low contrast between the foreground and background colors, as the GrabCut optimization algorithm is based on the probabilistic model for pixel color distribution.

2.5 Otsu's Thresholding

Thresholding involves classifying image pixels into categories of foreground and background pixels based on a given intensity threshold. Clearly, the main problem is to precisely detect the threshold which would give the optimal binarization result. Otsu's thresholding [5] selects the threshold value that minimizes the intra-class variance (or maximizes the inter-class variance), as a result of which the thresholding performs well only for images having bimodal intensity distribution.

2.6 Adaptive Thresholding

Fixed threshold cannot result in an appropriate segmentation when images are poorly illuminated. Local adaptive thresholding [6] solves this problem by selecting an individual threshold for each pixel based on intensity values of its neighbouring pixels. As a result, local adaptive thresholding performs much better even in the case of images for which global thresholding fails completely.

3 Experimental results

The performance of the above-listed six methods has been evaluated on the newly generated image dataset of 50 objects exposed to 4 different types of illumination conditions (the total of 200 images). Some of these images are shown in Fig. 2, while the illumination types are described below:
Type 1: Low intensity images having minimal contrast between the foreground and background colors (Fig. 2(a)).
Type 2: Images with sufficient lighting conditions having negligible shadows (Fig. 2(b)).
Type 3: Highly illuminated bright images with moderate shadows (Fig. 2(c)).
Type 4: Images with strong shadows having high intensity inhomogeneity (Fig. 2(d)).

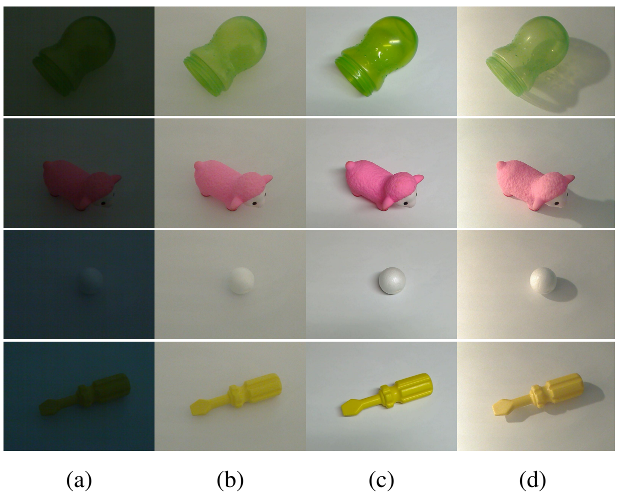


Figure 2: Examples of dataset images (a) Type 1 images (images of low intensity), (b) Type 2 images (images with no shadows), (c) Type 3 images (bright images with moderate shadows), (d) Type 4 images (images with strong shadows).

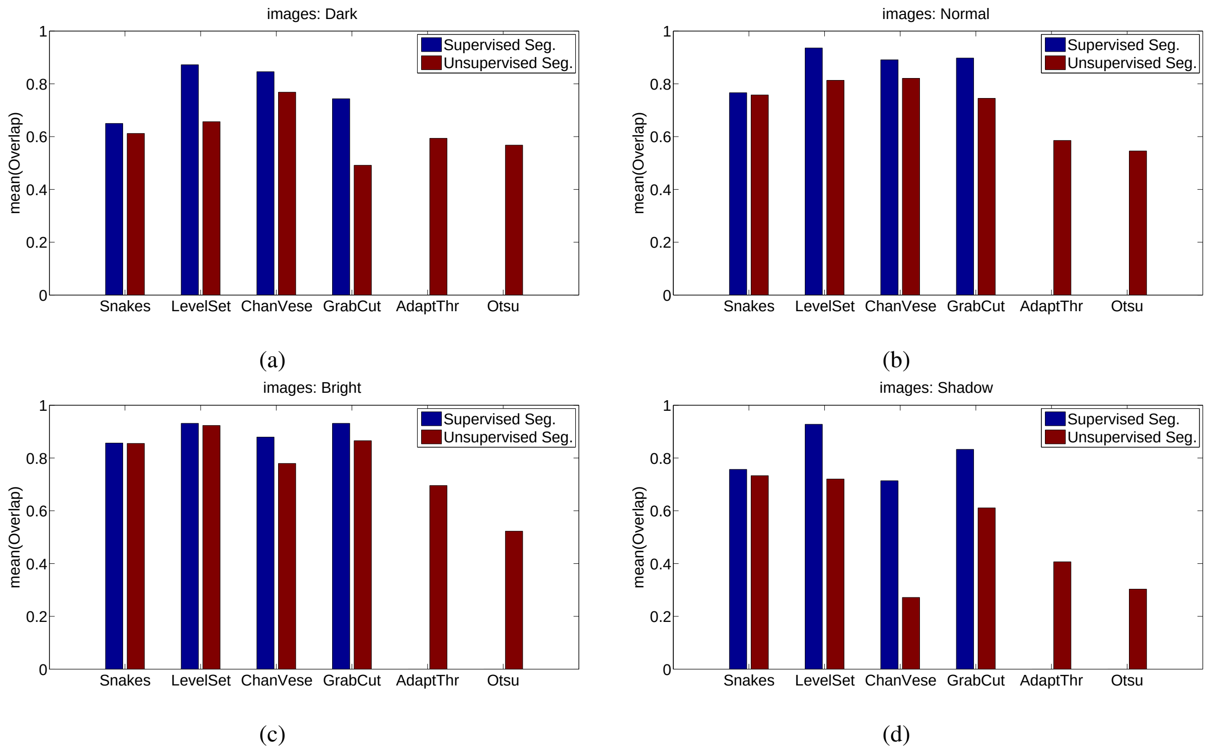


Figure 3: Comparison of segmentation methods (a) Type 1 images (low intensity images), (b) Type 2 images (images with no shadows), (c) Type 3 images (images with moderate shadows) , (d) Type 4 images (images with strong shadows).

All images were transformed to grayscale before applying the segmentation algorithms.

All the contour-based methods require an initial contour to be set. We used the apriori knowledge about images (one object in the center of a uniform background) to set the initial contour automatically. A rectangular contour (with size 10 percent less than the image size) was used to initialize Active contour, DRLSE and Chan Vese methods. GrabCut algorithm was initialized by declaring a square (20% of the image size) in the image center as foreground and all the pixels outside the rectangle (size 10% less than the image size) as background. We refer to this type of initialization as unsupervised, since we did not use the ground truth information for individual images. To verify the influence of the initialization to the segmentation process we also initialized the contours by considering the ground truth segmentations. This supervised initialization was done in the case of contour based methods by dilating the ground truth contour by 10 percent while the initialization for supervised GrabCut was done by declaring all the pixels inside the eroded ground truth as foreground and all pixels outside the dilated ground truth as background.

For all the methods we empirically determined the parameters that produce the best results, fixed the parameters and used them for segmenting all images.

For measuring the performance of the segmentation methods, the overlap measure was used, considering the segmented and ground truth regions as follows:

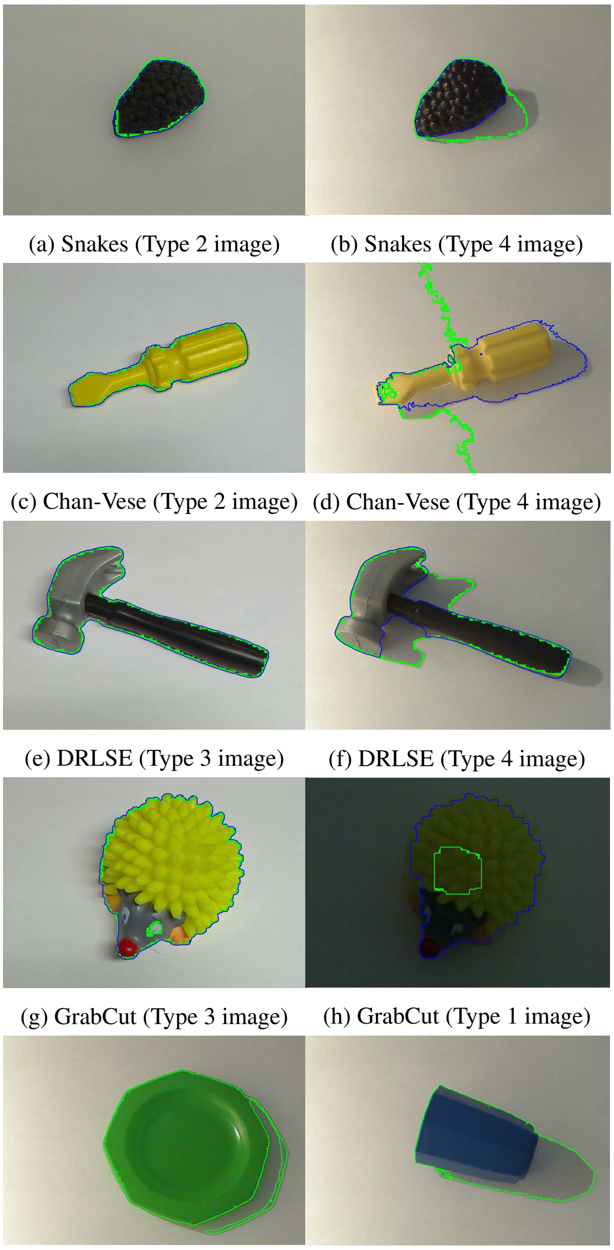
$$Overlap = \frac{|R_{result} \cap R_{groundTruth}|}{|R_{result} \cup R_{groundTruth}|} \quad . \qquad (1)$$

The results are depicted in Fig. 3 and are further analyzed in the remaining of this section.

Type 1 images: For images with low intensity and less contrast between the foreground and background pixels, the overlap accuracy of unsupervised Grabcut was found to be lower than with all other methods (Fig. 3(a)). This is because of the fact that GrabCut is based on the probabilistic model for pixel color distribution and this makes it ineffective to segment low-contrast images (an example is shown in Fig. 4(h)). The performance of unsupervised Chan Vese was found to be the best (almost similar to that of supervised GrabCut method) as the Chan Vese model is capable of segmenting images with minimal gradient. By providing perfect initialization the results significantly improve (by 50% for GrabCut and 30% for DRLSE) compared to its unsupervised counterpart.

Type 2 images: For images taken under sufficient lighting conditions with clear edges and good contrast all the methods performed well (Figs. 4(a) and (c)). Unsupervised DRLSE and Chan Vese model performed best with the mean overlap of 81% and 82% respectively, while in the case of the perfect supervised initialization DRLSE achieved even 94% of the mean overlap (Fig. 3(b)).

Type 3 images: For the images with moderate shadows, gradient based unsupervised contour methods (Snakes and DRLSE) performed surprisingly well with the results comparable to their supervised counterpart (Figs. 4(e) and (g)). This is because of the presence of good contrast (and hence high gradient) resulting from shadows along the object boundaries. Even though unsupervised DRLSE and GrabCut were able to achieve the mean over-



(a) Snakes (Type 2 image) (b) Snakes (Type 4 image)

(c) Chan-Vese (Type 2 image) (d) Chan-Vese (Type 4 image)

(e) DRLSE (Type 3 image) (f) DRLSE (Type 4 image)

(g) GrabCut (Type 3 image) (h) GrabCut (Type 1 image)

(i) Adap. Thr. (Type 4 image) (j) Otsu's Seg. (Type 4 image)

Figure 4: Performance of the segmentation methods on different images. Green contour is the result of unsupervised segmentation and blue contour is the result of supervised segmentation.

lap as high as 92% and 87% respectively (Fig. 3(c)), the performance of Chan Vese model decreased from Type 2 to Type 3 images. This is the consequence of the nonhomogeneous intensity distribution of the image. As the Chan Vese model is based on global image information, it fails to segment images with intensity inhomogeneities.

Type 4 images: Due to presence of strong shadows and nonuniform illumination, the performance of all the methods dropped by 25%, on average, from Type 3 to Type 4 images (Fig. 4 (b), (d), (f), (i) and (j)). The effect of intensity inhomogenity was significant on unsupervised Chan Vese model, with its accuracy reduced by 60% (Fig. 3(d)) .

4 Conclusion

In this paper, six segmentation methods were evaluated on the dataset of 200 images with varying lighting conditions. By analysing the results we can conclude that: **(a)** for images with low intensity and minimal gradient, Chan Vese method performs better than other unsupervised methods, **(b)** for images with no shadows and with uniform illumination, DRLSE and Chan Vese model outperform other methods, **(c)** for images with high gradient along edges and moderate shadows around the object, unsupervised edge-based active contour (Snakes and DRLSE) perform as good as their supervised counterpart, and **(d)** strong shadows and nonhomogeneous intensities significantly influence the effectiveness of the segmentation methods, especially the Chan Vese method.

The contour initialization turned out to be a very important step, so in our future work we will aim at improving the automatic initialization. We will also incorporate information about the color in the segmentation process. Furthermore, we will aim at finding an automatic way of setting parameters of individual methods to further improve the overlap accuracy of the segmentation algorithms.

References

[1] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *IJCV*, 1(4):321–331, 1988.

[2] T. F. Chan and L. A. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277, 2001.

[3] C. Li, C. Xu, C. Gui, and M. D. Fox. Distance regularized level set evolution and its application to image segmentation. *IEEE Transactions on Image Processing*, 19(12):3243–3254, 2010.

[4] C. Rother, V. Kolmogorov, and A. Blake. "GrabCut": Interactive foreground extraction using iterated graph cuts. *ACM Trans. Graph.*, 23(3):309–314, 2004.

[5] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man and Cybernetics*, 9(1):62–66, 1979.

[6] J. Sauvola and M. Pietikäinen. Adaptive document image binarization. *Pattern recognition*, 33:225–236, 2000.

[7] K. Zhang, H. Song, and L. Zhang. Active contours driven by local image fitting energy. *Pattern recognition*, 43(4):1199–1206, 2010.

[8] M. Jung, G. Peyre, and L. D. Cohen. Nonlocal active contours. *SIAM Journal on Imaging Sciences*, 5(3):1022–1054, 2012.

[9] S. Osher and J. A. Sethian. Fronts propagating with curvature-dependent speed: Algorithms based on hamilton-jacobi formulations. *Journal of Computational Physics*, 79(1):12 – 49, 1988.

[10] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 42(5):577–685, 1989.

[11] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE PAMI*, 23(11):1222–1239, 2001.