Pacific Graphics (2011)
Bing-Yu Chen, Jan Kautz, Tong-Yee Lee, and Ming C. Lin (Editors)

Short Papers

Fast Level Set Image Segmentation Using New Evolution Indicator Operators

Chunxia Xiao [1] and Jiajia Gan[1] and Xiangyun Hu[2]

[1] School of Computer, Wuhan University, Wuhan, China
[2] School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China

Abstract

We propose an effective level set evolution method for robust object segmentation in real images. We construct an effective region indicator and a multiscale edge indicator to adaptively guide the evolution of the level set function. The region indicator is built on the similarity map between image pixels and user specified interest regions, which is computed using Gaussian Mixture Models (GMM). To accurately detect object boundary, we propose to use a multi-scale edge indicator defined in the gradient domain of the multi-scale feature-preserving filtered image. Then, we develop a new mixing edge stop function based on these two indicators, which forces the level set to evolve adaptively based on the image content. Furthermore, we apply an acceleration approach to speed up our evolution process, which yields real time segmentation performance. As the results show, our approach is effective for image segmentation and works well to accurately detect the complex object boundaries in real-time.

Categories and Subject Descriptors (according to ACM CCS): I.4 [Computing methodologies]: Image Processing and Computer Vision—Applications

1. Introduction

Level set methods, first introduced by Osher and Sethian [OS88] for capturing moving fronts, have been widely used in image processing, computer vision, and computer graphics [Set00, OF03]. Level set methods work well due to its parameter-free representation, topological flexibility, effective numerical solution, and are especially useful in image segmentation, target tracking, image-based modeling, geometry processing, and physically-based fluid simulation [OF03]. Although level set methods have been successfully in active contour tracking, there are still many problems to be solved, which mainly focus on performance and effectiveness (see [LXGF05] for a survey).

In this paper, we propose a new level set image segmentation method aiming to segment objects in real natural images. The key idea is to construct an effective edge indicator and a region indicator to adaptively guide the evolution of level set function. The edge indicator is defined on multi-scale enhanced images which are immune from noises and texture details. The region indicator is built on the similarity map between the features on the propagating front and those at the user scribbles on the object. The similarity map is computed by probabilistic representation– multi-scale GMM (Gaussian Mixture Models). Furthermore, we combine these two indicators to develop a mixing evolution indicator, which forces the level set to evolve adaptively based on both the texture color distribution and gradient, so as to drive the motion of the contour toward the object boundaries adaptively. We also integrate the accelerated level set algorithm [SK05] into our system, to speed up the evolution process, which makes our system segment the image with moderate size in real-time and with little quality sacrificed.

User interactions can also be easily incorporated in our system. The whole process begins by user scribbling on regions of interest and background regions respectively, which can be considered as training data. Then we compute the similarity map between other image pixels and the training data using multi-scale GMM. This similarity map is especially effective in selecting appearance similar regions, not requiring the user to specify all the pixels to be segmented. In addition, it can preserve image structures. Our system can refine the segmentation as the user specifies more scribbles to segment the desired regions. This provides the user with full control of modification and refinement in each incremental step.

DOI: 10.2312/PE/PG/PG2011short/019-024

The proposed new level set formulation has following main advantages over traditional methods. First, using the proposed evolution indicators, our method works well on images with noise, weak boundaries, non-homogeneous regions,and texture details, to produce consistent segmentation results. Second, using the GMM based region indicator, the level set method is more robust to initial active contour setting. Third, the user interaction can be integrated into our method conveniently, which is especially useful in dealing with local boundary refinement. Finally, our method can be accelerated and can segment images with moderate size in real-time.

2. Background

The basic idea of level set methods [OS88] is to embed the moving interface $S \subset R^n$ as the zero isocontour of one higher-dimension implicit function $\phi : R^n \times R^+ \longrightarrow R$. In two spatial dimensions, the closed curve, denoted by C, can be implicitly represented as the zero level set $C(t) = \{(x,y)|\phi(x,y,t) = 0\}$ of a level set function $\phi(x,y,t)$. The speed function F in the evolution equation of the level set function is defined depending on the mean-curvature [OS88] or image edges information [CKS97].

In order to avoid drawbacks of re-initialization procedure in level set evolution, Li et al. [LXGF05] proposed a variational level set formulation without the need to re-initialization. This model accomplishes image segmentation by minimizing the following variational level set function:

$$E(\phi) = \mu P(\phi) + \lambda E_{edge}(g,\phi) + \nu E_{area}(g,\phi) \qquad (1)$$

where $\lambda > 0$ and ν are constants, the internal energy $P(\phi)$ penalizes the deviation of the level set function from a signed distance function during the evolution, which is defined by $P(\phi) = \int_\Omega \frac{1}{2}(|\nabla\phi| - 1)^2 dxdy$. The terms $E_{edge}(g,\phi)$ and $E_{area}(g,\phi)$, which drive the motion of zero level set of ϕ, are defined by $E_{edge}(g,\phi) = \int_\Omega g\delta(\phi)|\nabla\phi| dxdy$ and $E_{area}(g,\phi) = \int_\Omega gH(-\phi) dxdy$, respectively. The energy $E_{edge}(g,\phi)$ and $E_{area}(g,\phi)$ can be viewed as the weighted length of the zero level set $\partial\Omega$ and the weighted area of the region inside $\partial\Omega$, respectively. δ is the Dirac function, H is the Heaviside function, and g is the edge stop function of active contour, defined by

$$g = \frac{1}{1 + |\nabla G_\sigma * I|^2} \qquad (2)$$

where $\nabla G_\sigma * I$ is the gradient of Gaussian convolved image. During the evolution, the level set function ϕ is maintained close to a signed distance function, the zero level curve of ϕ moves toward the object boundaries, and stops at the boundaries when $E(\phi)$ reaches its minimum.

Although level set methods [OS88, LXGF05] work well for image segmentation, there are some limitations. Both methods [OS88, LXGF05], which use speed function F defined on mean-curvature [OS88] or use the stop function defined on image gradient [LXGF05], may produce undesirable results when processing real natural images with weak boundaries, or rich texture structures. The method [LXGF05] depends heavily on the initial contour selection, which determines the sign of the parameter ν. This limitation leads to inconvenient image segmentation. In this paper, we try to address above problems.

3. Description of the model

We define a novel energy function which incorporates edge-based information and region-based information simultaneously. We define a multi-scale edge indicater on the multi-scale edge-preserving smoothed images, which is more robust to image noises and texture details. We then define a region indicator, which is defined on the similarity map computed using Gaussian mixture model (GMM) to evolve level set function. We further combine these two indicators to develop a mixing edge stop function which corporately drive the level set curve to stop at the salient object boundaries. Finally, an acceleration approach is applied to our method to speed up the evolution process.

3.1. Multi-scale Edge Indicator

The noise and detail layers of an image may prevent edge stop function from producing desirable results. To get more robust and accurate segmentation, our edge stop function is built on the coarser images with detail layers attenuated. Since progressively coarser image levels increase the degree of abstraction in the resulting image [FFLS08] and produce more consistent similar regions, we build image gradient for each level and combine them in a spatially varying manner to produce a more consistent image gradient. Based on the combined gradient, we propose a multi-scale edge indicator for driving evolution of the active contour.

Let I be the input image for which we would like to construct M progressively coarser images. Utilizing an edge-preserving filter, for example bilateral filtering (BLF) [TM98, FAR07, FFLS08], we compute the progressively coarser version $I_1,\ldots,I_M$ of I. Then we compute the gradient on each coarser image I_j and define the multi-scale image gradient as the weighed gradient built from different levels:

$$Wgradients(I) = \frac{\sum_{j=1}^{M} w_j.\nabla I_j}{\sum_{j=1}^{M} w_j} \qquad (3)$$

where the weight w_j is defined as: $w_j = g_\sigma * (j \cdot e^{-\frac{1}{n}\Sigma|\nabla I_j|})$. n is the number of image pixels, the Gaussian convolution g_σ is used to locally smooth the weight. This weighted approach prefers coarser image by giving them larger weight because the color and luminance noises are smoothed more. In our experiments, we find out that just setting $M = 4$ is able to generate good results. The progressively coarser images can be computed in real time using bilateral grid methods [CPD07].

(a) (b) (c) (d) (e)

Figure 1: *(a) Initial contour, (b) Gaussian filtered image, (c) segmentation result driven by edge stop function defined on Gaussian filtered image (b) [LXGF05], (d) feature-preserved multi-scale coarsened image, (e) segmentation result by multi-scale edge indicator defined on (d).*

The edge indicator based on this multi-scale gradient is effective in identifying edges, which is defined as:

$$g_{Edge} = \frac{1}{1+|Wgradient(I)|^2} \tag{4}$$

Then, we introduce the edge energy functional based on the edge indicator:

$$E'_{edge}(g_{Edge}, \phi) = \int_\Omega g_{Edge}\delta(\phi)|\nabla\phi|\,dxdy \tag{5}$$

As illustrated in Fig. 1, by evolving level set function (1) using proposed multi-scale edge indicator, the segmentation result is better than the edge indicator defined on the gradient of the Gaussian filtered image. The detected boundaries are more accurate and smooth.

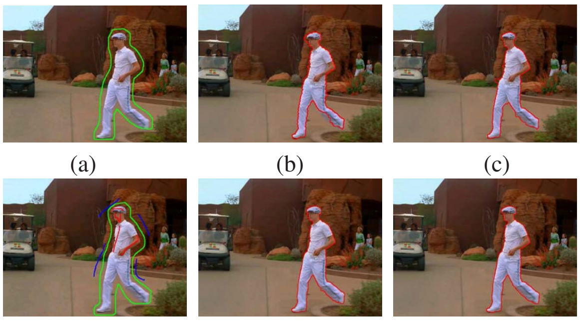
3.2. Region Indicator Based on GMM

Actually, most real images have rich color and texture information, when level set evolving via only gradient, it will lead to excessive evolution or premature convergence when dealing with images with weak edges, abundant texture structure and noises. As shown in Fig. 2, neither Gaussian filter nor multi-scale edge-preserving filter performs well if only gradient was considered. As an alternative, we drive the level set by the similarity map between the feature at the active contour and the user specified interest regions, which is estimated by the Gaussian mixture models (GMM).

GMM is a widely used model in image analysis for describing color distribution, the density function of GMM can be computed by the Expectation-Maximization algorithm.

We first select two sample sets from the foreground and background regions through user scribbles, as shown in Fig.2(d), then build the corresponding Gaussian mixture models GMM_f and GMM_b to represent the color PDF of these two sample sets. Let x_q represents the color feature vector of a given pixel q, based on Bayesian Decision Rule, we could judge whether the pixel q belongs to the foreground in the following way:

$$p(x_q|w_f)P(w_f) > p(x_q|w_b)P(w_b)$$



(a) (b) (c)

(d) (e) (f)

Figure 2: *(a) Initial contour, (b) results of [LXGF05], (c) results of our method using multi-scale edge-preserving filter, (d) initial contour, foreground data set, background data set, (e) results using edge indicator and region indicator, (f) results with mixed edge stopping function.*

where w_f and w_b are the forground class and background class,respectively. $P(w_f)$ and $P(w_b)$ are the prior probabilities, $p(x_q|w_f)$ and $p(x_q|w_b)$ are the likelihoods of w_f and w_b with respect to x_q, computed by GMM_f and GMM_b.

Let $g_f(q) = -In(p(x_q|w_f))$, $g_b(q) = -In(p(x_q|w_b))$ and $delt = In\frac{P(w_b)}{P(w_f)}$, then the above inequality is described as:

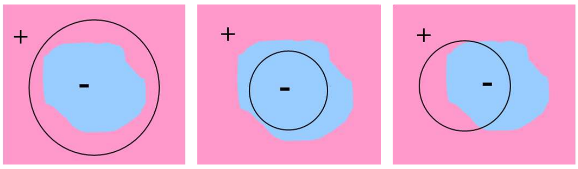
$$g_f(q) - g_b(q) + delt < 0 \tag{6}$$

The term $g_f(q) - g_b(q) + delt$ is called as the similarity map value in our paper.

To make the GMM estimation more robust to image noises and texture details, similar to multi-scale edge indicator, we define MultiGMM model as the weighed sum of GMMs built from different level coarser images:

$$MultiGMM(I) = \frac{\sum_{j=1}^{M} w_j.GMM(I_j)}{\sum_{j=1}^{M} w_j} \tag{7}$$

where the weight w_j is defined similar to Section 3.1. The similarity map value (6) is defined on the MultiGMM model instead of the original image intensity.



(a) (b) (c)

Figure 3: *The initial contour locates outside the object (a), inside the object (b) and has an arbitrary location (c). The blue region is the object and the red region is the background, the black curve is the initial contour. '+' and '-' represent the sign of the region indicator.*

Based on this similarity map, we introduce a region indicator, corresponding to each pixel q:

$$g_{Area}(q) = \left(g_f(q) - g_b(q) + delt\right) / \max f \quad (8)$$

where $maxf$ is the maxima value of $|g_f(q) - g_b(q) + delt|$. Then, a new region energy functional is defined as:

$$E'_{area}(g_{Area}, \phi) = \int_{\Omega} g_{Area} H(-\phi)\, dxdy \quad (9)$$

With the proposed multi-scale edge indicator and region indicator, a new level set energy functional is defined as:

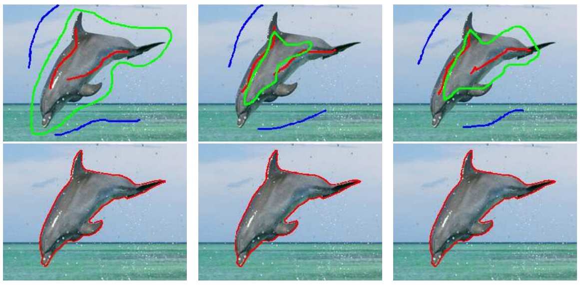
$$E^{egmm}(\phi) = \mu P(\phi) + \lambda E'_{edge} + E'_{area} \quad (10)$$

The external energy $\left(\lambda E'_{edge} + E'_{area}\right)$ drives the zero level set toward the object boundaries, where $\lambda > 0$ is a parameter controlling the effect of E'_{edge} and E'_{area}.

Being different from equation (1), the coefficient ν, which depends on the relative position of the initial contour to object of interest, is eliminated in our method (10). In equation(1), the coefficient ν should take positive value if the initial contours are placed outside the object and take negative value conversely. However, in our method, the area indicator g_{Area} itself contains the sign, as shown in Fig. 3. Therefore, there is no need to use the parameter ν to identify the relative location of the contours.

Fig. 3 shows the three possible location relations between the initial contour and the object. As proposed above, E'_{area} is the sum of g_{Area} at the pixels inside the zero level set contour. In these three cases, with evoluting, the positive components in E'_{area} decrease and the negative components increase, which leads to the gradually decreasing of E'_{area}. When the contour reaches the object boundary, E'_{area} takes its current minimum. Supposing the contour continues to move forward, the negative parts in E'_{area} will decrease and the positive parts will increase, which will inversely increase E'_{area}. Therefore, E'_{area} obtains its minimal at the boundary, which will lead to the termination of the evolution.

Fig. 2 shows the segmentation results applying method [LXGF05] (Eq. 1) and our method (Eq. 10), which is obviously improved. The parameters $\mu = 0.03$ and $\lambda = 3.2$ in our method. In Fig. 6 and Fig.4, we give an example that our method can detect the desirable object, even if the initial contour is not accurately selected.



(a) (b) (c)

Figure 4: *The initial contour locates outside the object (a), inside the object (b) and has an arbitrary location (c).*

3.3. Mixing stop function

GMM region indicator works well at the regions which obviously belong to the foreground or the background distribution, and edge indicator works at the regions with abrupt intensity change. Thus, we use both of them to speed up the evolution. The basic idea is that, for the regions distinctively belonging to the foreground or the background, we mainly use g_{Area} to drive the motion of the curve, while for the color overlapping regions, g_{Edge} is used to further evolve the curve.

Based on the analysis above, to speed up curve evolution, we propose the following mixed region indicator $g_{MixArea}$, which is the weighted average of g_{Area} and sg_{Edge}:

$$g_{MixArea} = \omega_A \cdot g_{Area} + \omega_E \cdot sg_{Edge} \quad (11)$$

where sg_{Edge} is a signed edge indicator, whose sign is in according with the sign of g_{Area} at pixel q, that is:

$$sg_{Edge}(q) = \begin{cases} (-1) \cdot g_{Edge}(q) & if\ g_{Area}(q) < 0 \\ g_{Edge}(q) & otherwise \end{cases}$$

The way to set the weights ω_A and ω_E is that, giving a pixel larger weight ω_A if it corresponds to relative larger $|g_{Area}|$, and weight ω_E is defined in a similar way. Let avg_A and avg_E be the average values of $|g_{Area}|$ and $|sg_{Edge}|$ of all pixels in the image, respectively, and $m = g_{Area}/avg_A$, $n = sg_{Edge}/avg_E$, then the weights ω_A and ω_E are defined as: $\omega_A = m/(m+n)$, $\omega_E = n/(m+n)$.

We modify the region energy functional as follows:

$$E''_{area}(g_{MixArea}, \phi) = \int_{\Omega} g_{MixArea} H(-\phi)\, dxdy \quad (12)$$

then a new level set energy functional is defined as:

$$E^{EMixgmm}(\phi) = \mu P(\phi) + \lambda E'_{edge} + E''_{area} \quad (13)$$

The ultimate evolution equation of the level set function is:

$$\frac{\partial\phi}{\partial t} = \mu\left[\Delta\phi - div\left(\frac{\nabla\phi}{|\nabla\phi|}\right)\right] + \lambda\delta(\phi)\,div\left(g_{Edge}\frac{\nabla\phi}{|\nabla\phi|}\right) + g_{MixArea}\delta(\phi) \quad (14)$$

Fig. 2 shows the results applying the mixed region indicator, compared with methods using the edge indicator and region indicator independently, the segmentation is more accurate. The parameter $\mu = 0.03$ and $\lambda = 3.2$ in the Eq.(14).

3.4. Fast level set method

High computational cost of level set limits its application, especially for segmenting high resolution images. In this section, inspired by the method [SK05], we accelerate our method to accomplish interactive image segmentation.

For the evolving curve C, we define two lists of neighboring pixels *Lout* and *Lin* as:

$$Lout = \{q|\phi(q) > 0 \text{ and } \exists p \in N_4(q) \text{ such that } \phi(p) < 0\}$$
$$Lin = \{q|\phi(q) < 0 \text{ and } \exists p \in N_4(q) \text{ such that } \phi(p) > 0\}$$

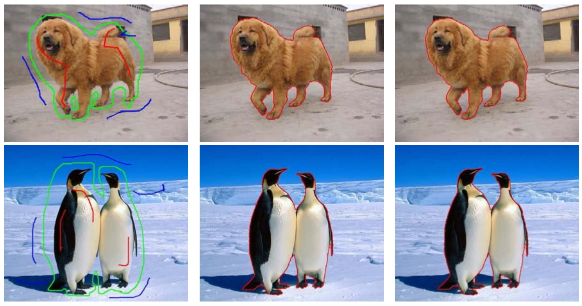
Where $N_4(q)$ is a 4-connected neighborhood of a pixel q with q itself removed. By switching the neighboring pixels between the two lists *Lout* and *Lin*, the curve can be moved inward or outward for one pixel everywhere in one scan. Then we use a Gaussian filtering process to further speed up the algorithm. The whole process is implemented by two cycle operations.

We set the speed function F_d proposed in [SK05] as our mixing region indicator. In cycle one, we let $F_d = -g_{MixArea}$, when scanning each pixel in *Lout* and *Lin*, we use both F_d and g_{Edge} to identify whether a pixel should be switched in or out. For each pixel q in *Lout*, if $F_d(q) > 0$ and $g_{Edge}(q) > threshold$, we switch it in, in the same manner, for each pixel q in *Lin*, if $F_d(q) < 0$ and $g_{Edge}(q) > threshold$, we switch it out. Restrained by the edge indicator, the evolving curve can be driven more accurately around the boundaries. In cycle two, we smooth the level set function value at each pixel in *Lout* and *Lin* by a Gaussian filter to further switch pixels in or out. Applying this accelerating approach, the curve can be evolved at pixel resolution.

Fig.5 shows the segmentation results using and without using acceleration. In these experiments, it averagely takes $0.083s$ and $11.094s$ using and without using acceleration respectively, to segment a 588×501 image, which shows considerable improvement. In addition, with the same user-interaction, the segmentation results are visually similar, which shows the effectiveness and efficiency of the acceleration operation.

4. The experimental results and discussion

To validate the effectiveness of the proposed approach, we test it using different real images.We compare our algorithm with the popular level set segmentation methods and other interactive image segmentation. All algorithms are implemented using C++ on a laptop with an Intel Core Duo 2.6GHz CPU and 2G memory.



(a) (b) (c)

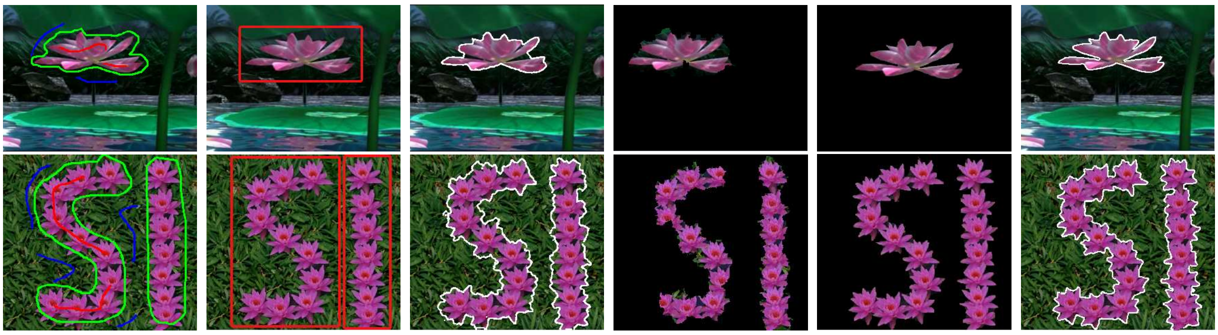
Figure 5: *Segmentation results before and after acceleration. (a) initial contour, foreground and background scribes, (b) our results without acceleration. (c) our results with acceleration.*

In Fig.6, we compare with the methods [QWH06], Lazy snapping [LSTS04] and Grabcut [RKB04]. Qu et al. [QWH06] evolves the level set based on the pattern-continuity of underlying image, and the evolution stops at the boundary where the pattern exhibits abrupt change, it can not work well when processing the images with rich texture and weak edges, as shown in Fig.6. Lazy Snapping [LSTS04] applies graph cuts to create binary masks, which prefers to generate more closed and consistent regions. We perform Lazy snapping method using the code presented on Webside [†]. Both Grabcut [RKB04] and our method perform very well. The advantage of Grabcut is its simple initial interaction (marked rectangle), however, it usually needs more iteration optimization and interactions to receive final satisfied segmentation, which makes Grabcut not efficient. Given the initial scribbles, our method can simultaneously select multi objects that are similar to the stroked regions and perform in real-time with fewer following interactions. It takes Grabcut 1.1s to segment the image with size 321×242, while our method takes 0.07s.

To measure the accuracy of our method, we use the ground truth database in Grabcut [RKB04] for reference, the error rates are $2.15 \pm 0.15\%$ for our method compared with $2.13 \pm 0.18\%$ for GrabCut.

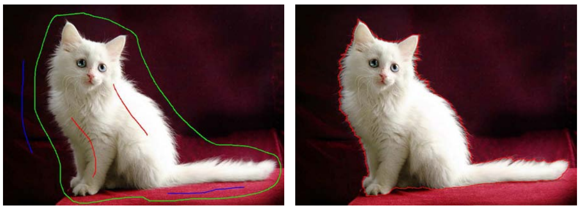
limitations: Our method can not achieve satisfied segmentation results when the objects we are interested in are very complex, such as containing fur or hair. As illustrated in Fig. 7, our method can not effectively segment the object out. To solve this problem, we need to incorporate the image matting techniques [WC07] to further refine our method.

† http://www.cs.cmu.edu/ mohitg/segmentation.htm



(a) (b) (c) (d) (e) (f)

Figure 6: *Image segmentation results comparison. (a)interaction in our method, (b) interaction in [RKB04], (c) results of [QWH06], (d) result of Lazy snapping [LSTS04], (e) result of Grabcut [RKB04], (f) our results.*



(a) (b)

Figure 7: *Failure results. (a) initial contour, foreground and background scribes, (b) segmentation results.*

5. Conclusion and future work

In this paper, we mainly aim to address the segmentation of images with rich color textures and week boundaries. The key idea is constructing an effective edge indicator and a region indicator to drive the evolution of the level set function. In addition, an acceleration algorithm has also been integrated into our method to obtain real-time image segmentation. The experimental results indicate that the proposed method can not only effectively segment the object out but also obtain smooth boundaries.

In the future, we will apply image matting techniques to our methods to further refine the level set based image segmentation. When the foreground and background regions have very similar color distribution, we would like to incorporate the low-level cues such as image saliency into the level set formulation to improve the segmentation results.

Acknowledgment

We would like to thank the anonymous reviewers for their valuable comments and insightful suggestions. This work was partly supported by NSFC (No. 60803081, No. 61070081).

References

[CKS97] Caselles V., Kimmel R., Sapiro G.: Geodesic active contours. *International journal of computer vision 22*, 1 (1997), 61–79.

[CPD07] Chen J., Paris S., Durand F.: Real-time edge-aware image processing with the bilateral grid. *ACM Transactions on Graphics (TOG) 26*, 3 (2007), 103.

[FAR07] Fattal R., Agrawala M., Rusinkiewicz S.: Multiscale shape and detail enhancement from multi-light image collections. *ACM Transactions on Graphics 26*, 3 (2007), 51.

[FFLS08] Farbman Z., Fattal R., Lischinski D., Szeliski R.: Edge-preserving decompositions for multi-scale tone and detail manipulation. *ACM Trans. Graph 27*, 3 (2008), 67.

[LSTS04] Li Y., Sun J., Tang C., Shum H.: Lazy snapping. *ACM Transactions on Graphics (ToG) 23*, 3 (2004), 303–308.

[LXGF05] Li C., Xu C., Gui C., Fox M.: Level set evolution without re-initialization: A new variational formulation. *CVPR* (2005), 1063–6919.

[OF03] Osher S., Fedkiw R.: *Level set methods and dynamic implicit surfaces*. Springer Verlag, 2003.

[OS88] Osher S., Sethian J.: Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations. *Journal of computational physics 79*, 1 (1988), 12–49.

[QWH06] Qu Y., Wong T., Heng P.: Manga colorization. *ACM Transactions on Graphics (TOG) 25*, 3 (2006), 1214–1220.

[RKB04] Rother C., Kolmogorov V., Blake A.: Grabcut: Interactive foreground extraction using iterated graph cuts. In *ACM Transactions on Graphics (TOG)* (2004), vol. 23, ACM, pp. 309–314.

[Set00] Sethian J.: *Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science*. Cambridge University Press, 2000.

[SK05] Shi Y., Karl W.: Real-time tracking using level sets. *CVPR* (2005), 34–41.

[TM98] Tomasi C., Manduchi R.: Bilateral filtering for gray and color images. In *ICCV* (1998), IEEE, pp. 839–846.

[WC07] Wang J., Cohen M.: Image and video matting: a survey. *Foundations and Trends in Computer Graphics and Vision 3*, 2 (2007), 97–175.