ROBUST INTERACTIVE IMAGE SEGMENTATION WITH WEAK SUPERVISION FOR MOBILE TOUCH SCREEN DEVICES

Tinghuai Wang
Nokia Technologies
Finland
tinghuai.wang@nokia.com

Huiling Wang
Lappeenranta University of Technology
Finland
huiling.wang@lut.fi

Lixin Fan
Nokia Technologies
Finland
lixin.fan@nokia.com

ABSTRACT

In this paper, we present a robust and efficient approach for segmenting images with less and intuitive user interaction, particularly targeted for mobile touch screen devices. Our approach combines geodesic distance information with the flexibility of level set methods in energy minimization, leveraging the complementary strengths of each to promote accurate boundary placement and strong region connectivity while requiring less user interaction. To maximize the user-provided prior knowledge, we further propose a weakly supervised seed generation algorithm which enables image object segmentation without user-provided background seeds. Our approach provides a practical solution for visual object cutout on mobile touch screen devices, facilitating various media manipulation applications. We describe such a use case to selectively create oil painting effects on images. We demonstrate that our approach is less sensitive to seed placement and better at edge localization, whilst requiring less user interaction, compared with the state-of-the-art methods.

ABSTRACT

In this paper, we present a robust and efficient approach for segmenting images with less and intuitive user interaction, particularly targeted for mobile touch screen devices. Our approach combines geodesic distance information with the flexibility of level set methods in energy minimization, leveraging the complementary strengths of each to promote accurate boundary placement and strong region connectivity while requiring less user interaction. To maximize the user-provided prior knowledge, we further propose a weakly supervised seed generation algorithm which enables image object segmentation without user-provided background seeds. Our approach provides a practical solution for visual object cutout on mobile touch screen devices, facilitating various media manipulation applications. We describe such a use case to selectively create oil painting effects on images. We demonstrate that our approach is less sensitive to seed placement and better at edge localization, whilst requiring less user interaction, compared with the state-of-the-art methods.

1. INTRODUCTION

Interactive image segmentation is becoming growingly popular especially on mobile touch screen devices to facilitate spatially localized media manipulation, since prior knowledge about the desired object and background can be easily defined with simple user interactions such as marking of object boundaries [1, 2], placing a bounding box around the foreground object [3, 4], and loosely drawing scribbles on foreground/background regions [5, 6, 7]. Among these different interaction modes, drawing scribbles is of the most interests for touch screen devices since it provides sufficiently rich information about the desired foreground and background regions while requiring less and flexible user interventions. Despite of the significant advances delivered in recent years, two open issues prevent the scribble-based approach from being used by massive mobile device end-users. First, to draw scribbles on both the foreground and background objects is too troublesome for compact touch screens. The need to switch between foreground and background scribbles also complicates the UI. Second, it is cumbersome for novice end users to perform substantial fine-tunings to correct mis-segmentations, especially the noisy boundaries and disjoint regions which may severely affect the quality of the target applications.

To tackle these two compelling issues, we propose in this paper a novel interactive segmentation method that a) relieves the burden of drawing scribbles on background regions; b) automatically promotes accurate boundary placements and fills contiguous, coherent regions without substantial user corrections. For the first goal, we formulate seed generation as a weakly supervised segment annotation problem, initially generating background seeds from the given image border, pruning mis-labelled background seeds depending on their *superpixel geodesic distance* from foreground seeds and generating new background seeds via cascaded classification (see Sec. 3). For the second goal, we proposes a geodesic level set framework that combines *geodesic region information* with the flexibility of *the level set methods* in energy minimization.

1.1. Related Work

Geodesic distance has been used for interactive image segmentation driven by scribbles [6], which selectively fill the desired region by expanding from the interior of the selected object outwards without explicitly considering the object boundary. This make it advantageous for segmenting objects with complex topologies, whilst it may suffer from a bias that favors shorter paths from the seeds. These methods may also fail to accurately identify the real object boundaries

due to the lack of an explicit optimization framework taking into account of edge contrast, something at which level set or graph cut based methods generally excel.

Despite the improvements reported by the approaches combining geodesic distance information and graph cut [8, 9], these approaches are limited by the inherent bias of graph cut towards shorter paths, i.e. small segments as the optimization sums over the boundaries of segmented regions. Anh *et al.* [10] proposed a continuous-domain convex active contour model combining geodesic distance based probability with color distance based probability, which essentially relies on the difference between geodesic distances from FG/BG seeds. This difference generally accounts for the sensitivity to seeds placement and disjointed regions in geodesic segmentation.

Level set methods [11] neatly enable the minimization of energy functionals. Caselles *et al.* [12] proposed the edge-based *geodesic active contour model* for image segmentation as a geometric alternative for snakes, to obtain the optimal image partition by finding the set of minimal length geodesic curves that are attracted by the real region boundaries. More robust approaches that encode region information were proposed later by Paragios *et al.* [13]. Higher level prior knowledge such as geometric shape priors have also been introduced to level set framework [14, 15].

2. GEODESIC LEVEL SET

The general idea of *level set* methods is that a contour $\mathcal{C}$ in a domain Ω can be represented by the zero level set of a higher level embedding function ϕ: $\Omega \rightarrow \Re$. Evolving the contour $\mathcal{C}$ is achieved by evolving the embedding function ϕ which is defined as the signed distance function with $\phi > 0$ inside the contour, $\phi < 0$ outside the contour and $|\nabla\phi| = 1$ almost everywhere.

The evolution of the level set function ϕ is governed by the the Euler-Lagrange equation which minimizes $E(\phi)$ at $\frac{\partial\phi}{\partial t} = -\frac{\partial E(\phi)}{\partial\phi}$. These methods are known as variational level set methods [11]. Thus the segmentation boundary $\mathcal{C}$ is derived by obtaining the optimal ϕ at the zero level as $\mathcal{C} = \{x \in \Omega \mid \phi(x) = 0\}$.

Following the level set paradigm, we propose a new energy functional to incorporate geodesic region information:

$$E(\phi) = \alpha E_g(\phi) + E_e(\phi) \tag{1}$$

where α is an importance coefficient, $E_g(\phi)$ is the proposed *geodesic region term* and $E_e(\phi)$, as defined in [16], is the edge term which consists of two components. The first component of edge term is the geodesic active contour component which is minimized when the zero level contour of ϕ is located at the object boundaries; it is essential for the contour evolution to stop at the desired object boundary. The second component is the ballooning component which speeds up the motion of the zero level contour in the level set evolution process to avoid short-cutting. We incorporate this ballooning component mainly to address the short-cutting problem, for which a local weighting strategy has been adopted within the graph cut framework (e.g. see [9]). The reader is referred to [16] for details of the definition of edge term $E_e(\phi)$. We introduce the derivation of the proposed geodesic region term below.

In order to reduce the sensitivity to seeds placement, we propose a novel *geodesic region term* $E_g(\phi)$ which measures the statistics of the geodesic distance of all pixels instead of the individual geodesic distance used in [6]. Following Paragios *et al.* [13], an optimal partition $\hat{\mathcal{P}}(\Omega)$ of the image plane Ω can be computed by maximizing the *a posterior* probability $p(\mathcal{P}(\Omega)|I)$ for the given image I. Applying Bayes' rule, it can be expressed as $p(\mathcal{P}(\Omega)|I) \propto p(I|\mathcal{P}(\Omega))p(\mathcal{P}(\Omega))$. Under the given prior $p(\mathcal{P}(\Omega))$, optimal two-region partition is achieved by maximizing $p(I|\mathcal{P}(\Omega)) = p(I|\Omega^+)p(I|\Omega^-)$, where $\Omega^+ = \{x|\phi(\mathbf{x}) > 0\}$ and $\Omega^- = \{x|\phi(\mathbf{x}) < 0\}$ represent the regions inside and outside the contour in question, and $p(I|\Omega^+)$, $p(I|\Omega^-)$ are likelihoods of foreground/background regions respectively. We derive $E_g(\phi)$ as

$$E_g(\phi) = -[\log p(I|\Omega^+) + \log p(I|\Omega^-)]. \tag{2}$$

The minimization of this energy function is equivalent to the maximization of posterior probability of an optimal partition $\hat{\mathcal{P}}(\Omega)$.

We assume that the image I in each region is characterized by the individual relative geodesic distance $G_l(x|\theta)$ ($l \in \{\mathcal{F}, \mathcal{B}\}$) at different locations x and $G_l(x|\theta)$ values are i.i.d. We reduce (2) to $E_g(\phi) = -\int_\Omega (H(\phi)\log G_\mathcal{F}(x|\theta) + (1-H(\phi))\log G_\mathcal{B}(x|\theta)d\mathbf{x}$, where θ represents the foreground and background color models, H is the Heaviside function, and $G_l(x|\theta) = 1 - \frac{D_l(x)}{D_\mathcal{F}(x)+D_\mathcal{B}(x)}$ which represents the relative geodesic distance from current pixel x to label l based on color models θ.

Let $\Omega_\mathcal{F}$ be the foreground seed and $\Omega_\mathcal{B}$ be the background seed. The geodesic distance from each of the two sets of seeds for every pixel x is computed as $D_l(x) = \min_{s\in\Omega_l} d(s, x), l \in \{\mathcal{F}, \mathcal{B}\}$ where

$$d_l(s_1, s_2) := \min_{L_{s_1,s_2}} \int_0^1 |W_l(L_{s_1,s_2}(p)) \cdot \dot{L}_{s_1,s_2}(p)|dp \tag{3}$$

$L_{s_1,s_2}(p)$ is a path parameterized by $p = [0, 1]$ connecting the pixels s_1 to s_2 respectively, and $W_l(s)$ gives the geodesic weight. We set $W_l(s)$ as the gradient of the likelihood that a pixel belongs to label l as in [6], i.e., $W_l(s) = \nabla P_l(C(s))$, where $P_l(c) = \frac{Pr(c|l)}{Pr(c|\mathcal{F})+Pr(c|\mathcal{B})}$, $C(s)$ is the color at s and $Pr(c|l)$ the likelihood that a pixel with color c belongs to label l. The foreground and background color models are represented by Gaussian Mixture Model (GMM) learned from observations of foreground and background seeds respectively.

This proposed geodesic region term provides region-based information which encodes both color distribution and spatial information. The incorporation of geodesic information is crucial to alleviate the tendency of either under- or over-segmentation in feature distance based models, e.g. [13], when the foreground and background appearance models are indistinct. Embedding geodesic information into a Bayesian inference framework, our proposed geodesic region term significantly reduces the sensitivity to seeds placement and the number of disjoint regions.

Finally, we use the standard gradient descent method to minimize the energy functional (1) as $\frac{\partial\phi}{\partial t} = -\frac{\partial E_g(\phi)}{\partial\phi} - \frac{\partial E_e(\phi)}{\partial\phi}$, where the gradient flow of proposed geodesic region

term is deducted as follows:

$$\frac{\partial E_g(\phi)}{\partial \phi} = \delta(\phi) \log \frac{G_{\mathcal{B}}(x)}{G_{\mathcal{F}}(x)}$$

where δ is the Dirac delta function.

To substantially reduce the computational cost of level set method, we adopt the narrow band method [17] to confine the computation to a narrow band around the zero level set contour. In our implementation, the embedding function ϕ is initialized by extracting the contour of the user-provided foreground brush stroke. The embedding function is assigned as 2 inside the contour and -2 outside the contour. We empirically choose the parameter in the formulation to be $\alpha = 2.0$, by optimizing performance against ground truth over a training set of 50 images. Despite the availability of adaptive weighting methods [9], our setting proved to be versatile for a wide variety of images without increasing computational complexity.

3. SEED GENERATION VIA WEAKLY SUPERVISED LEARNING

Scribble-based segmentation approaches enable intuitive user interfaces for many image editing applications on desktop PCs e.g. [6, 2]. For compact touch screens of mobile devices, it is important to allow end-users to specify their object of interests by simple finger gestures e.g. tapping or sweeping. Existing scribble-based segmentation methods, which request user-specified seeds for both the foreground and background regions, are too troublesome for this purpose — not only switching between foreground and background scribbles complicates the UI, but also the sensible choice of the location and amount of background scribbles to draw also troubles mobile device end-users. Therefore, we propose a weakly supervised learning driven seed generation algorithm to automatically generate background seeds from user-provided foreground seeds. An example of the seed generation process is shown in Fig. 1.

We start by over-segmenting the image into small coherent regions i.e. superpixels, using the fast graph-based segmentation method in which each segment is characterized by a CIE Lab color histogram [18] (Fig. 1 (a)). The hand-drawn scribbles are associated with those segments that intersect with scribbles. This set of user-selected foreground segments form the initial set of foreground seeds $\Omega_{\mathcal{F}}$. The set of initial background seeds $\Omega_{\mathcal{B}}$ consists of the segments along the border of image plane (Fig. 1 (c)). Note that the set of initial background seeds are chosen based on the assumption that foreground objects rarely intersect with image border. Even though this assumption may be violated under certain circumstances (note the area marked by the ellipse in Fig. 1 (c)), a further label pruning step is adopted to eliminate mis-labelled segments (Fig. 1 (d)). More background seeds (areas marked by the ellipses in Fig. 1 (e))) are generated by performing classification in a cascaded setting to classify the rest unlabelled segments. Details of seed generation are elaborated below.

The seed generation is formulated as a weakly supervised segment annotation problem. The initial foreground seeds $\Omega_{\mathcal{F}}$ form the set (denoted as $\mathcal{P}$) of all instances with a positive label whilst the initial background seeds $\Omega_{\mathcal{B}}$ form the set of all negative instances denoted as $\mathcal{N}$. This initial set of seeds provides a pool of weakly labeled data $\{< s_1, l_1 >, \ldots, < s_N, l_N >\}$, where s_i is segment i, and $l_i \in \{-1, 1\}$ is the label for segment i, with the label being positive if the segment $s_i \in \Omega_{\mathcal{F}}$, and negative if $s_i \in \Omega_{\mathcal{B}}$. Since our positive data was weakly labeled by user input, we can assume that the segments labeled as positive are (with rare exceptions) correctly labeled whilst the negative instances $\mathcal{N}$ may contain label noise. Our task then is to trim and generate negative segments. We propose a two-phase algorithm to i) prune the false negative instances $\mathcal{N}_{fn}$ in $\mathcal{N}$ and ii) perform cascaded classification to propagate the negative instances.

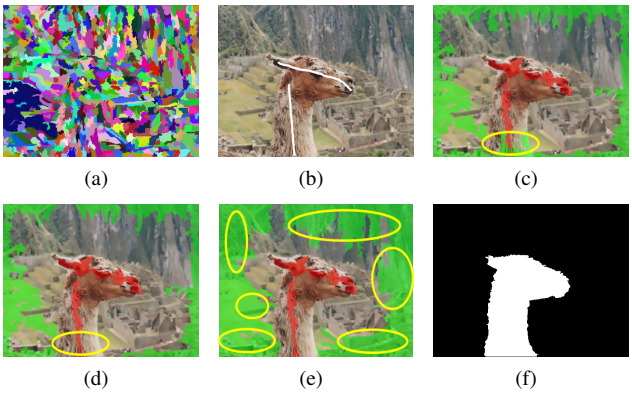

Fig. 1. Seed generation. (a) superpixel map (b) user-provided FG scribbles applies on underlying segments (c) initial FG (red) and BG (green) seeds (d) FG (red) and BG (green) seeds after label noise pruning in Sec. 3.1 (d) seeds after seed generation in Sec. 3.2. (e) segmentation using generated seeds in (d). pruning of label noise and generated BG seeds are highlighted in ellipses.

3.1. Label Noise Pruning

We formulate the label noise pruning as a problem of ranking the elements of $\mathcal{N}$ in decreasing order of a score, $S(s_i)$ such that top-ranked elements correspond to $\mathcal{N}_{tn}$; thresholding at a particular rank prunes $\mathcal{N}_{fn}$. Appearance based distance measure alone is typically not discriminative when the foreground and background are not visually distinct or the appearance of the background segments are diverse. It is natural to encode spatial information into the appearance distance measure, where geodesic distance could again be adopted.

Distinctively, we propose to compute the weighted distance functions (geodesics) at the superpixel level which could take the advantage of segment-based measure instead of using the noisy pixel-based measure or the inaccurate color models learned from the available seeds with label noise. We set the geodesic weight $W(s_1, s_2)$ in (3) to be the gradient of segment features. Specifically, we compute the χ^2 distance between the features of each pair of consecutive segments along the path, i.e. CIE Lab color histograms h_{s_1} and h_{s_2}, to approximate $W(s_1, s_2)$. We define $W(s_1, s_2) =$

$\frac{1}{2}\sum_{k=1}^{\mathcal{K}} \frac{[h_{s_1}(k)-h_{s_2}(k)]^2}{h_{s_1}(k)+h_{s_2}(k)}$, where $\mathcal{K}$ is the total number of bins present in the descriptor (=20 for each CIE Lab channel). L_{s_1,s_2} in (3) is defined as the Euclidean distance between the geometrical centers of consecutive segments. Applying these new definitions in (3), we can compute the *segment geodesic distance* from each instance in $\mathcal{N}$ to $\mathcal{P}$. Intuitively, true negative segments are identified as those among $\mathcal{N}$ whose nearest neighbor among $\mathcal{P}$ is as far as possible. To this end we rank all elements in $\mathcal{N}$ in descending order of their distances from $\mathcal{P}$ and keep the top-ranked (e.g. 60% percentile) segments as true negatives while pruning the rest false negatives. Note that this segment geodesic distance is the key to the success of the weakly supervised segment annotation.

3.2. Seed Generation

To propagate the negative instances, we propose a cascade classification process which can be structured into two stages: (i) supervised discriminative learning to generate a candidate pool of negative instances based on appearance similarity (ii) geodesic distance based selection to enforce spatial constraint.

Given the labeled instances in $\mathcal{P}$ and $\mathcal{N}$, we adopt linear SVM to learn a discriminative classifier to approach segment annotation, propagating true negative instance labels to unlabeled segments. Specifically, we train a linear SVM on the CIE Lab color histograms of labeled segments, which is then applied on the unlabelled segments. Those segments that are classified negative form a candidate pool $\mathcal{C}$ of negative instances.

Since Phase I results in a reliable set of $\mathcal{N}$ through label noise pruning, the superpixel geodesic distance measure can also be computed from $\mathcal{N}$ to every superpixel candidate $r_i \in \mathcal{C}$. Let $D_{\mathcal{N}}(r_i)$ and $D_{\mathcal{P}}(r_i)$ be the superpixel geodesic distances from $\mathcal{N}$ and $\mathcal{P}$ to candidate superpixel r_i respectively. The relative geodesic distance $G_{\mathcal{P}}(r_i)$ from $\mathcal{P}$ to r_i can be computed as $G_{\mathcal{P}}(r_i) = \frac{D_{\mathcal{P}}(r_i)}{D_{\mathcal{N}}(r_i)+D_{\mathcal{P}}(r_i)}$. Superpixels $r_i \in \mathcal{C}$ which satisfy $G_{\mathcal{P}}(r_i) > 0.5$ are added to $\mathcal{N}$ as negative instances, i.e. new background seeds. By doing so, the number of background seeds is effectively increased yet without bringing in noisy seeds near foreground instances.

4. EXPERIMENTS AND COMPARISONS

In this section, we evaluate the accuracy of the geodesic level set framework and the effectiveness of the proposed seed generation algorithm, in comparison with six state of the art scribble-based segmentation methods, namely, the *geodesic active regions (LS)* [13], *geodesic segmentation (GD)* [6], *regular graph cut (GC)* [5], *geodesic graph cut (GDGC)* [9], *star-convexity prior graph cut (SSGC)* [7], and its variant using geodesic information (*GSGC*) [8] [1]. Two schemes of the proposed geodesic level set framework are compared in experiments, i.e. the geodesic level set approach without using the seed generation algorithm (*GDLS*) and the one augmented with seed generation (*GDLS-SG*). Segmentation accuracy of all methods are compared both quantitatively and qualitatively, and results are elaborated in following subsections.

We have compared all segmentation methods on a dataset consisting of both the GrabCut dataset (49 images) [3] and 74 images from the Berkeley $BSDS500$ dataset [19]. The GrabCut dataset comes with ground truth for the task of interactive segmentation. We manually labelled the ground truth of the rest of the images for evaluation. For a fair comparison of performances, segmentations of different methods are driven by the same set of user-provided scribbles except that background seeds are provided by the proposed seed generation algorithm for our *GDLS-SG* method.

4.1. Quantitative Evaluation

We adopt the *segmentation overlap score (SOS)* used in the VOC segmentation challenge [20] to quantify segmentation accuracy against a manual specified ground-truth: $SOS = \frac{\mathbf{y}\cap\mathbf{y}_{gt}}{\mathbf{y}\cup\mathbf{y}_{gt}}$, where $\mathbf{y}$ denotes the set of segmented foreground pixels and $\mathbf{y}_{gt}$ the ground truth foreground pixels. For each method, SOS is measured for each test image and statistics of SOS for all test images is used to indicate the accuracy of each method.

Table 4 summarizes the results of this quantitative comparison, which shows that the proposed *GDLS* achieves the highest average segmentation overlap score (91.10%) and outperforms all other state-of-the-art methods by varying margins between 1% and 8%. It is worth noting that, the proposed approach *GDLS* delivers a performance gain of 2.7% over *LS* and 7.9% over *GD* respectively. Noticeably, the *GDLS* method also yields the tightest standard deviation (0.0674) across the whole test dataset. The low standard deviation confirms the superior *robustness* of the proposed *GDLS* method across various natural images in the dataset, which is in accordance with theoretic analysis of the adopted energy terms defined in Section 2.

The geodesic level set method enhanced by the proposed seed generation algorithm (*GDLS-SG*) is able to yield comparable performance (90.90%) of *GDLS* with minor degradation less than 1%. Even though no user-provided scribbles are used for *GDLS-SG*, it still outperforms other state-of-the-art methods e.g. the geodesic active regions framework (88.36%) and the geodesic segmentation (83.18%). Since the *GDLS-SG* method significantly reduces user interaction efforts for compact touch screen devices, it is actually an appealing and preferred solution in practice.

All graph cut variants enhanced by spatial or shape constraint outperform the regular graph cut method which reaches a low average overlap score 84.80%. By taking advantage of its robust foreground and background prior, *GSGC* (90.26%) outperforms *SSGC* (89.97%) and *GDGC* (89.66%). Table 4 also illustrates a noticeable trend that the last three methods (*GDGC*, *SSGC*, *GSGC*) yield more robust segmentation as indicated by the tighter standard deviation compared to regular graph cut.

[1] We used the publicly available source code from: `ttp://www.robots.ox.ac.uk/~vgg/researc/iseg/`

	GDLS	GDLS-SG	LS	GD	GC	GDGC	SSGC	GSGC
Average (%)	**91.10**	90.90	88.36	83.18	84.80	89.66	89.97	90.26
Median (%)	**92.56**	92.45	90.54	87.12	86.53	91.44	92.09	92.27
Std	**0.0674**	0.0689	0.0859	0.1274	0.1137	0.0756	0.0704	0.0691

Table 1. Statistics of segmentation overlap scores from our objective comparison on the dataset

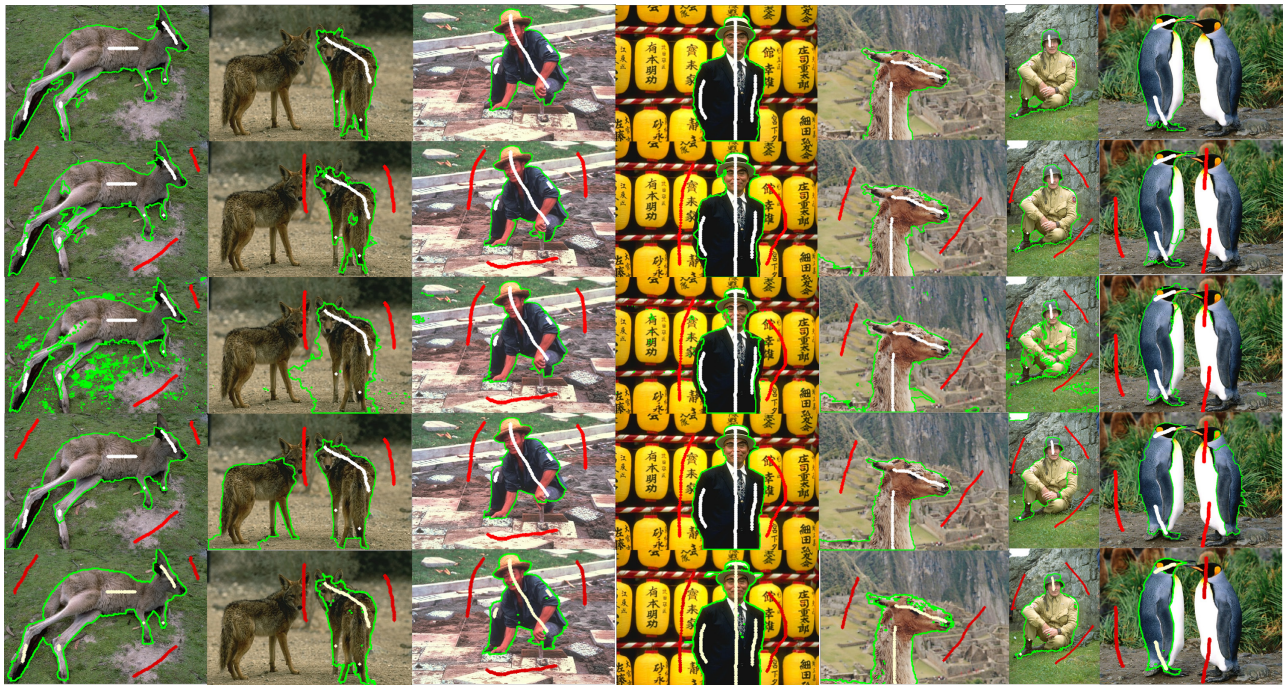

Fig. 2. Comparison of proposed geodesic level set framework with seed generation (rows 1) with LS (row 2) [13], GD (row 3) [6], GDGC (row 4) [9], and GSGC (row 5) [8].

4.2. Qualitative Evaluation and Use Case

We perform qualitative comparison through visual inspection for all test images and some typical results are summarized in Fig. 2. Specifically, we selectively compare the proposed *GDLS-SG* with *LS*, as well as methods incorporating geodesic information, i.e. *GD*, *GDGC*, and *GSGC* in Fig. 2 to demonstrate the benefit of combining geodesic region information and level set. The proposed *GDLS-SG* method demonstrates superior segmentation accuracy in challenging scenarios, e.g. when the foreground and background exhibit similar colors (columns 2, 5 and 7), the background is very cluttered or complex (columns 3 and 4), or the foreground objects have complex topology (column 1). These segmentation results exhibit accurate boundary placement and strong region connectivity.

As a representative level set based method which combines both the geodesic active contour model and region information, *LS* exhibits diversified performance (row 2 in Fig. 2). On the one hand, it fills the desired region with less disjoint regions by expanding from the interior of the selected object outwards and respecting the object boundary. On the other hand, it tends to either under-segment or over-segment objects when the foreground and background color models are indistinct (see columns 1, 4 and 5 for mis-segmentation examples). More mis-segmentations can be observed in results from *GD* (row 3), due to the lack of explicit or implicit edge information and the sensitivity to seeds placement of the method. In case of complex scenes, geodesic segmentation fails to distinguish the foreground from background given limited user-provided seeds (e.g. column 2). By exploiting the geodesic distance information, *GDGC* (row 4) effectively avoids some short cutting problems and results in less disjoint regions. Nevertheless, it tends to degenerate into regular graph cut when the foreground and background objects have similar color (see e.g. the mis-segmented wolf in columns 2 and 7), due to the dependency of the adaptive weighting on the classification performance of the color models. *GSGC* (row 5) extends star-convexity prior introduced in *SSGC* by replacing the straight lines (Euclidean rays) with geodesic paths. However, this strong connectivity constraint tends to be sensitive to underlying color models learned from seeds examples. Consequently, under-segmentation are often observed for regions with similar color distribution in the close vicinity of foreground (columns 2 and 4), as it tends to push the contour outwards along the geodesic paths. The shortcomings discussed above are addressed, to various extents, in the proposed geodesic level set framework.

As an application of our proposed algorithm on mobile

Fig. 3. Segmentation (L) and background SBR effects (R).

devices, we apply an automatic stroke-based painterly rendering (SBR) algorithm by Shugrina *et al.* [21] to create virtual brush strokes for oil painting and photo composition as shown in Fig. 3, where the background scene is stylized as oil painting effects. The whole process only takes a couple of user scribbles as shown in Fig. 2 to indicate the object of interest. As to the computational complexity, the running time of the proposed segmentation method on a commodity PC is ~ 0.5 seconds per VGA image (640×480).

5. CONCLUSION

This paper has presented a robust and efficient image segmentation method, in which both geodesic region information and geodesic active contour model are combined in a level set framework. Substantial performance gain from the proposed approach has been achieved over the individual level set method without geodesic information or geodesic segmentation methods. Under this framework, we have also proposed a novel weakly supervised seed generation algorithm which relieves the burden of providing background scribbles which is crucial to reduce interaction efforts on compact touch screen devices.

6. REFERENCES

[1] A. Blake, C. Rother, M. Brown, P. Perez, and P. Torr, "Interactive image segmentation using an adaptive gmmrf model," in *ECCV*, 2004, pp. 428–441.

[2] J. Wang, M. Agrawala, and M. F. Cohen, "Soft scissors: an interactive tool for realtime high quality matting," in *SIGGRAPH*. 2007, pp. 585–594, ACM.

[3] C. Rother, V. Kolmogorov, and A. Blake, "Grabcut - interactive foreground extraction using iterated graph cuts," in *SIGGRAPH*. 2004, ACM.

[4] V. Lempitsky, P. Kohli, C. Rother, and T. Sharp, "Image segmentation with a bounding box prior," in *ICCV*, 2009, pp. 277–284.

[5] Y. Boykov and M.-P. Jolly, "Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images," in *ICCV*, 2001, pp. 105–112.

[6] X. Bai and G. Sapiro, "A geodesic framework for fast interactive image and video segmentation and matting," in *ICCV*, 2007, pp. 1–8.

[7] Olga Veksler, "Star shape prior for graph-cut image segmentation," in *ECCV*, 2008, pp. 454–467.

[8] V. Gulshan, C. Rother, A. Criminisi, A. Blake, and A. Zisserman, "Geodesic star convexity for interactive image segmentation," in *CVPR*, 2010, pp. 3129–3136.

[9] B. L. Price, B. S. Morse, and S. Cohen, "Geodesic graph cut for interactive image segmentation," in *CVPR*, 2010, pp. 3161–3168.

[10] N. Anh, J. Cai, J. Zhang, and J. Zheng, "Robust interactive image segmentation using convex active contours," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3734–3743, 2012.

[11] H. Zhao, T. Chan, B. Merriman, and S. Osher, "A variational level set approach to multiphase motion," *Journal of Computational Physics*, pp. 179–195, 1996.

[12] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," in *ICCV*, 1995, pp. 694–699.

[13] N. Paragios and R. Deriche, "Geodesic active regions for supervised texture segmentation," in *ICCV*, 1999, pp. 926–932.

[14] M. Rousson and N. Paragios, "Shape priors for level set representations," in *ECCV*. 2002, pp. 78–92, Springer.

[15] T. Wang, B. Han, and J. Collomosse, "Touchcut: Fast image and video segmentation using single-touch interaction," *CVIU*, vol. 120, no. 3, pp. 14–30, 2014.

[16] C. Li, C. Xu, C. Gui, and M. D. Fox, "Distance regularized level set evolution and its application to image segmentation," *IEEE Trans. Image Process.*, vol. 19, no. 12, pp. 3243–3254, 2010.

[17] D. Adalsteinsson and J. Sethian, "A fast level set method for propagating interfaces," *Journal of Computational Physics*, pp. 269–277, 1995.

[18] P. Felzenszwalb and D. Huttenlocher, "Efficient graph-based image segmentation," *IJCV*, vol. 59, no. 2, pp. 167–181, 2004.

[19] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 5, pp. 898–916, May 2011.

[20] M. Everingham, L. J. Van Gool, C. K. I. Williams, J. M. Winn, and A. Zisserman, "The pascal visual object classes (voc) challenge," *IJCV*, vol. 88, no. 2, pp. 303–338, 2010.

[21] M. Shugrina, M. Betke, and J. Collomosse, "Empathic painting: Interactive stylization using observed emotional state," in *Proc. NPAR*, 2006, pp. 87–96.