

Winding Number Constrained Contour Detection

Yansheng Ming, *Student Member, IEEE*, Hongdong Li, *Member, IEEE*, and Xuming He, *Member, IEEE*

Abstract—Salient contour detection can benefit from integration of both contour cues and region cues. However, this task is difficult due to different nature of region representation and contour representation. To solve this problem, this paper proposes an energy minimization framework based on winding number constraints. In this framework, both region cues such as color/texture homogeneity and contour cues such as local contrast and continuity are represented in a joint objective function which has both region and contour labels. The key problem is how to design constraints which ensure the topological consistency of the two kinds of labels. Our technique is based on the topological concept of winding number. Using a fast method for winding number computation, a small number of linear constraints is derived to ensure label consistency. Our method is instantiated by ratio-based energy functions. By successfully integrating both region and contour cues, our method shows advantages over competitive methods. Our method is extended to incorporate user interaction, which lead to further improvements.

Index Terms—contour detection, image segmentation, winding number

I. INTRODUCTION

OBJECT contours play an important role in image understanding. They not only localize the support region of objects, but also reveal their shape which is a strong cue for recognition. Therefore, extracting a small set of clean and semantically meaningful contours may simplify subsequent high-level image understanding tasks [48]. However, salient contour extraction is a challenging task. First of all, the problem is highly under-constrained. Without any knowledge of objects, there are few prior constraints on the number of contours and their shape. Secondly, various factors such as texture, shadow, and lack of contrast could undermine the accuracy of contour detection.

To deal with these difficulties, a robust strategy is to use multiple types of cues. Many contour detection methods have been developed based on different cues, e.g. figure/background contrast [42], [50], [9], contour smoothness [39], closure [3], [30] and symmetry [43]. These cues often call for different representations. Region cues, such as the figure/background contrast, are more conveniently represented in 2D region segmentation methods [42], [50], [9], whereas contour cues such as smoothness are better represented in edge domain. For example, methods such as active contour method [19] focus on 1D contour detection.

Over the years, there has been a trend of jointly using contour cues and region cues for object contour detection. They can be roughly categorized as continuous or discrete

methods. In continuous domain, the active contour model was adapted to use both region and contour cues [45] [53]. In discrete domain, contour cues (such as curvature) have been introduced to region segmentation methods (e.g. [44], *intervening contour* approach in [23]). However, their techniques for cue integration are often tailored to their specific problems e.g. relying on heuristic or an unduly complicated model [55], and lack explicit interaction between region and contour labels.

This paper attempts a more general, yet efficient approach that tightly integrates both region cues and contour cues. We consider the contour extraction problem within an energy minimization framework. Our objective function is designed to encode various region and contour cues by explicitly introducing both 2D region labels and 1D contour labels. Then we use a set of constraints to ensure the consistency of region and contour labels. Compared to previous methods, this framework allows for higher flexibility in choosing energy functions.

The key to this framework is the design of constraints, which should have the following three properties. First of all, these constraints should ensure the topological correctness of solutions. For example, regions with different labels should be separated by contours, and object boundaries should not be fragmented. Second, these constraints should not be too restrictive such that there is no feasible solution. Last but not least, constraints should be encoded efficiently, and thus induce minimal computational costs, e.g., preferably a small number of linear constraints. However, it is recognized that, constraints satisfying these conditions are not easy to design. Recent work [3] argued that in order to ensure the closedness condition, exponentially many constraints are needed.

In this work, we propose a novel and simple method to describe the region-contour consistency relationship, based on the classic “winding number” concept from the field of algebraic topology [29]. By definition, a winding number, which involves several closed (but not necessarily simple) planar curves and a point in the plane, refers to the number of times the curves revolve around this point. Our key technique is to restrict region labels to be the winding numbers of contours, so that a small number of linear constraints can provably ensure region/contour consistency.

Our approach can be considered as an example of the “duality” relationship between region and contour. Being dual, the property of one can be converted into the property of the other. A well-established example is the application of Green’s theorem in the plane. Draw a simple Jordan curve (i.e. closed and non-self-intersecting contour) in the plane. Many quantities (such as its areas) defined on the 2D region can be computed efficiently via 1D line integral along the contour.

We apply our framework to the scenario in which regions have binary labels, and focus on the foreground contour detection. Using the winding number technique, we integrate

Y. Ming is with the College of Engineering and Computer Science, Australian National University, Australia, e-mail: yansheng.ming@anu.edu.au.

H. Li is with ANU and NICTA, e-mail: hongdong.li@anu.edu.au

X. He is with NICTA and ANU, e-mail: xuming.he@nicta.com.au

The material in this paper appeared in part at the 2013 IEEE conference on pattern recognition and computer vision.

region segmentation cues into the ratio-based contour detection formulation [49], [44]. Our objective function includes three components for contour saliency, region similarity and contour smoothness, respectively, and can be efficiently solved by linear programming.

Our method is evaluated on several public datasets, and is compared with pure contour or region based approaches. Our method is also extended to incorporate user interaction for interactive image segmentation. Although this paper focuses on middle-level perceptual grouping, we believe that our method is applicable to high-level tasks where both contour cues and region cues are helpful (e.g. object detection).

The paper is organized as follows. Section II discusses relevant methods in the literature. Section III introduces the winding number concept. Section IV presents our general energy minimization framework for cue combination. In Section V, this framework is instantiated by ratio-based energy functions. Experimental results by these energy functions are shown in Section VI, and conclusion is in Section VII. Compared to our conference paper [31], this paper has many new materials. First, our method is compared with more methods on more datasets, and we include additional results for visual comparisons. Second, our method is extended to incorporate user interaction for the purpose of interactive segmentation. Third, several illustrations for key concepts are also added.

II. RELATED WORK

Our work is closely related to contour detection and image segmentation methods. Contour detection methods aim to extract one dimensional curves which have physical meaning in an image. For example, these curves may correspond to object boundaries or surface markings. Human labeling, such as those in [5], is often used as ground truth for benchmarking.

Most of contour grouping methods start with local edge detection. Early work applies filters of special design to an image and estimates edge location from filter responses [10], [36]. Later, learning based methods obtain more robust performance. e.g. [27], [37], [21], due to their ability for using multiple cues such as texture, brightness, and color. However, ambiguities in local information prevent local detectors from obtaining clean results. A recent study [57] shows that some best local edge detectors have comparable accuracy with human, suggesting that the superiority of human contour detection lies in high-level image understanding.

Many methods have employed mid-level cues, often inspired by Gestalt laws of grouping [34]. For example, the Gestalt law of good continuity is popularly employed as a prominent contour cue [26], [52], [39] [20], [49]. In addition, the contour closure prior, based on a psychological study [22], has been addressed in [3], [30]. The global Pb method, an integration of several successful techniques, has state of the art performance on the BSDS dataset [5]. In addition to mid-level grouping cues, category-specific information is also used for detecting semantic contours [56], [16]. Interested readers are referred to [35] for a comprehensive review.

Image segmentation methods aim to partition image pixels into coherent and meaningful regions. For example, figure-ground segmentation methods separate whole objects from the

background. Supervised segmentation methods can segment an image into regions and recognize the category of each region [17]. Not using supervised information, our method belongs to the category of unsupervised segmentation methods including active contours [19], level set [41] and watershed method [47]. We pay special attention to methods belonging to graph partition framework. By representing image pixels as coupled graph nodes, image segmentation problem is transformed into a graph partition problem. Notable methods under this framework include the max-flow method [9], normalized cuts [42] and the ratio cut [50]. Moreover, spectral K-means [6], Felzenszwalb et al.'s graph-based method [13], SWA method [2] can also be considered as instances of this framework. In addition, the idea of graph partition also underlies interactive segmentation methods such as [38].

Some work has attempted to incorporate both region cues and contour cues. Intervening contour [23] is one of the early efforts to use local contour saliency for region segmentation. This method is built into a state of the art segmentation algorithm [5]. Yu et al. [54] incorporate edge information in their Markov random field. Tabb and Ahuja integrate both cues for low-level structure detection [46]. GPAC method [45] is one of the best variational method for cue integration in continuous domain. However their gradient descend inference method is susceptible to local minimal. Our method, however, solves linear programs for globally optimal solutions. Stahl and Wang [44] modify the ratio contour method [49] in order to extract more regular shape. The same ratio objective function is implemented by superpixels in [25], and improved results are obtained. In Nicolls and Torr's work [33], edge labels are determined by a linear transform of region labels, based on discrete topology of planar graph. However, each edge in their method has a predetermined orientation. For interactive segmentation, [40] uses local constraints to achieve boundary-region consistency. In contrast, the winding number constraints used by our method are global constraints. The winding number concept not only leads to a smaller number of constraints but also provides a clearer understanding about region contour interaction. Winding numbers have recently been used for extracting volume representation from 3D meshes [18].

Winding number is one of the basic concept in topology [32] [14]. Unfortunately, it is often confused with another topological concept, rotation index, which is defined as the total rotation angle of tangent if one travels along a curve [51]. Rotation index is often called rotation number, even winding number in the literature [28]. In computer vision field, rotation index has been used for ensuring contour topology in [12]. To the best of our knowledge, winding number has not been applied in perceptual grouping contexts.

III. WINDING NUMBER AND ITS FAST COMPUTATION

In this work, contours are considered as curves in two dimensional Euclidean plane E^2 . A curve is defined as the image of a continuous map from an interval of real line to the plane. A curve is closed if its end point coincides with the start point. A closed curve is also called a loop. The

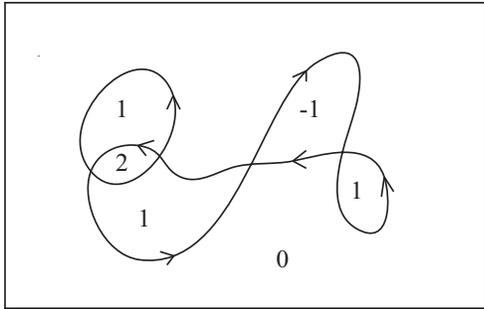


Fig. 1. Winding numbers induced by closed contours.

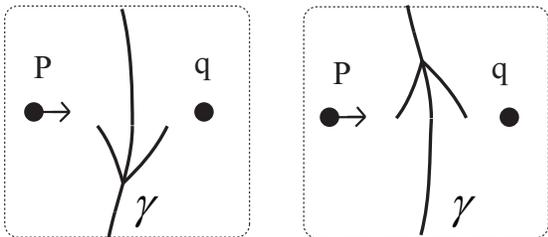


Fig. 2. Left: $wn_p = wn_q - 1$. Right: $wn_p = wn_q + 1$. wn_p and wn_q are the winding numbers of point p and q respectively.

winding number of a loop about a point is defined as the number of times the loop travels around the point in counter-clockwise direction. Here, the loop does not need to be *simple*, i.e. it is allowed to intersect with itself. As shown in [14], any connected component of the plane has a constant and integral winding number. Therefore, we can assign a single winding number to a connected region. We further extend winding number concept to a plane with a finite number of curves. The winding number of a region is defined as the sum of the winding numbers of all loops. Figure 1 shows the winding numbers of different regions induced by two closed curves. Take the region labeled with a winding number 2 for example, both two curves travel around this region in counter-clockwise direction, and each curve induces a winding number 1 to the region.

If winding numbers were to be computed according to the definition, we have to perform integration along loops in the plane. However, there is a fast method to compute winding numbers using the *crossing rule* [32]: “If a loop is moving from our left to our right [our right to our left] as we cross it, its winding number around us increases [decreases] by one.”

Let wn_p denote the winding number of point p , and y_α denote the binary label of edge α indicating whether the edge is active. The edge extraction step is explained in Section IV-A. An illustration of the crossing rule is shown in Figure 2.

According to the crossing rule, if we draw an path from point p to point q , the winding number difference is

$$wn_p - wn_q = \sum_{\alpha \in L_{pq}} y_\alpha - \sum_{\beta \in R_{pq}} y_\beta \quad (1)$$

where L_{pq} and R_{pq} denote the indices of edges crossing from right to left, and edges crossing from left to right on

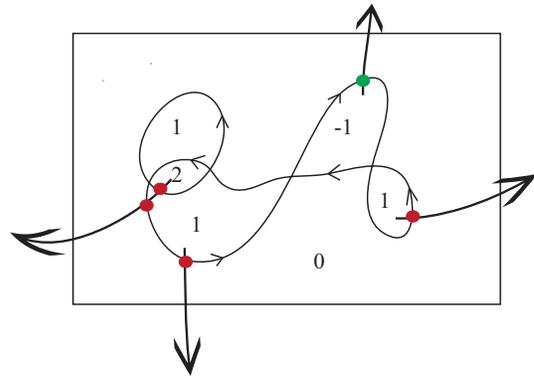


Fig. 3. Fast winding number computation. Draw an arbitrary path to outside of the image frame, the winding number of a point equals the number of edges crossing from right (red dot) minus the number of edges crossing from left (green dot).

that path, respectively. Please note that winding numbers are solely determined by the loops, it follows that Eq (1) holds regardless of the choice of paths. The intuition behind can be compared to climbing a hill. Whichever path we may choose, the total number of steps upward minus the total number of steps downward should be determined by the height of the hill. Please note that a winding number is invariant to the path used in calculation.

If we predetermine the winding number of some points in the plane to zero, e.g. the points very far away from any curve. Then we draw an arbitrary path starting from the inside of the region p to that point, then we can apply Eq (1) to calculate the winding number of the region as follows:

$$wn_p = \sum_{\alpha \in L_p} y_\alpha - \sum_{\beta \in R_p} y_\beta \quad (2)$$

where L_p and R_p are the edges crossing from right to left, and edges crossing from left to right, respectively. This method is illustrated in Figure 3.

This fast computation method (2) will be used to connect region labels to edge labels in our energy minimization framework. This property will make winding number calculation robust in practice.

IV. REGION-BOUNDARY CONSISTENT CONTOUR EXTRACTION

This section presents our method for region-boundary consistent contour extraction. Section IV-A discusses how to extract basic edge and region hypotheses from an image. Section IV-B discusses our general energy minimization framework and its consistency condition. Section IV-C shows how to use winding numbers to enforce this condition in a simplified way.

A. Basic edge and region hypotheses

Salient contours are those which human can see from noisy background. Please note that they are not necessarily part of salient objects which draw immediate human attention. We formulate the salient contour extraction problem as an

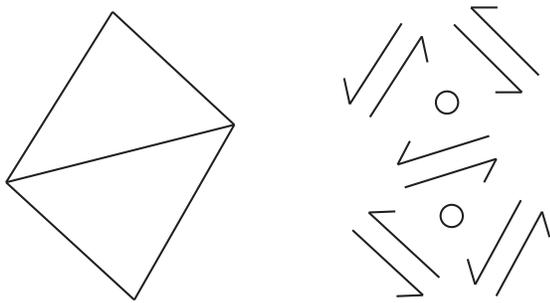


Fig. 4. Examples of region and edge hypotheses. The left shows two triangular image regions and their edges. The right shows the edge and region hypotheses extracted from the image. Two circles denote the variables of two regions, and each arrow represents a variable of a directed edge.

energy minimization problem defined on both region and edge hypotheses. We choose superpixel over-segmentation [24] as a means to provide sufficient edge and region hypotheses. Each superpixel provides an atom region hypothesis. The boundary of each superpixel are linearized into a number of edge-elements. For each element, two oppositely directed (bi-directional) edge hypotheses called conjugate edges are introduced. It is important to note that our winding number formulation is not restricted to the superpixel setup, but applies to general boundary-region graphs as well. An illustration of this setup is shown in Figure 4. These two triangles give rise to two atom regions and eight directed edges.

Let the variable $\mathbf{x} = \{x_i | i = 1 \dots N_r\}$ denote the labels of N_r atom regions, and $\mathbf{y} = \{y_j | j = 1 \dots N_e\}$ denote the binary labels of N_e edges. The edge label space is denoted as $\mathcal{Y} = \{0, 1\}^{N_e}$. The label space of all region variables is denoted as \mathcal{X} . Although the winding number concept is potentially applicable for multiple-label segmentation problems, our work focuses on the binary-label figure/ground segmentation problem. In other words, we let $\mathcal{X} = \{0, 1\}^{N_r}$.

B. Energy function and consistency condition

The problem of salient contour extraction is considered in the energy minimization framework, as many previous work did [19], [44], [21], [42]. The energy function $E(\mathbf{x}, \mathbf{y})$ represents various cues, such as figure-background contrast and contour smoothness, depending on applications. Our main contribution is a set of constraints which ensure the topological correctness of solutions. Together, the basic energy-minimization problem has the following form:

$$\min_{\mathbf{x}, \mathbf{y}} E(\mathbf{x}, \mathbf{y}) \quad (3)$$

$$s.t. \quad \Phi_W(\mathbf{x}, \mathbf{y}) = 0 \quad (4)$$

$$\Phi_C(\mathbf{x}, \mathbf{y}) = 0 \quad (5)$$

$$\mathbf{x} \in \mathcal{X}, \mathbf{y} \in \mathcal{Y} \quad (6)$$

In order to guarantee the topological validity of labels, two sets of constraints are set up. The constraint set Φ_C is the edge continuity constraints, as follows:

$$\sum_{i \in j_{in}} y_i = \sum_{i \in j_{out}} y_i, \quad \forall j \in V, \quad (7)$$

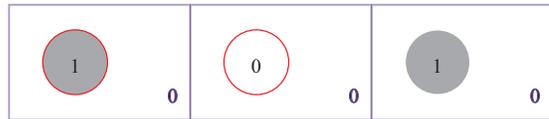


Fig. 5. The left image shows a consistent region and edge configurations. The middle figure is not consistent because two regions separated by the contour have the same labels. The right figure is not consistent because two adjacent regions with different labels are not separated by contour.

where j is from the vertex index set V ; j_{in} and j_{out} denote edges indices heading into and moving out of the vertex j , respectively. These constraints say that the net flow at every vertex is zero. For a flow network without source and sink, all the flows can be decomposed into a set of cycles. Therefore, our method aims to extract a set of closed curves as contours.

The constraints Φ_W denote the winding number constraints which ensure the consistency of region and contour labels. The specific consistency condition used in this paper is that:

If an edge is active, its adjacent (i.e. incident) regions must have different region labels; if two adjacent regions have different labels, one of the edge elements in-between must be active.

This condition guarantees that every edge must be part of a region boundary, and every region is enclosed contours. Figure 5 shows one correct labeling and two incorrect cases that violate the condition. At first glance, this condition does not have anything to do with winding numbers. Instead, it can be formulated as follows:

$$|y_m - y_n| = \mathbf{1}(x_i - x_j = 0), \quad \forall (i, j) \in G, \quad (8)$$

where x_i and x_j denote the labels of two adjacent regions, and G denotes the set of indices of adjacent regions. Variables y_m and y_n denote two conjugate edges separating these two regions. The function $\mathbf{1}(\cdot)$ equals one if its argument is true, and equals zero otherwise. Although constraints (8) are sufficient for the consistency condition, they are expensive to implement due to their non-linearity. Even if the energy function E is convex, the whole energy minimization problem will generally turn out to be non-convex with these non-linear constraints. Next section will show that this condition is guaranteed by linear constraints based on the winding number concept.

C. Winding number constraints

We realize that the winding number concept, from topological study, provides an elegant and effective means to parameterize the region-contour consistency condition for image segmentation. We have reached the following key step called *winding number technique*:

The label of a region can be identified by its winding number induced by contours.

The winding numbers can be used as region labels due to properties discussed in Section III. For example, region labels need to be integral and have the same value in connected regions. These conditions are proved to be true for winding

numbers. More importantly, the winding number constraints guarantee the consistency between region labels and contour labels. First of all, due to Eq (9), the winding numbers of adjacent regions will be different if one of the conjugate edges between them is active.¹ Secondly, Eq (9) also suggests that two regions which are not separated by any edges must have the same winding number. In other words, these two regions cannot have different labels. Thirdly, the existence of winding numbers also guarantee the feasibility of our problem (3). In conclusion, this winding number scheme does encode the region-contour consistency condition.

The benefit of such winding number scheme also lies in that: it leads to a small number of linear constraints. This can be made evident by examining the fast computation procedure of winding number computation in Eq (2). To adapt it to our problem, we assume that every image is enclosed by a rectangular border, region outside of the image has a label zero. Therefore, the winding number of region i is computed as:

$$x_i = \sum_{\alpha \in P_i} y_\alpha - \sum_{\beta \in N_i} y_\beta, \quad \forall i \quad (9)$$

where P_i and N_i are the edges crossing from right to left, and edges crossing from left to right, respectively. Eq (9) for all atom regions together can be represented as the following winding number constraint, denoted as Φ_W in Eq (4):

$$\mathbf{x} = M\mathbf{y}, \quad (10)$$

where M is a matrix whose entries are 0, 1, or -1 . Take the i -th row of M for example, $M(i, \alpha) = 1, \forall \alpha \in P_i$; $M(i, \beta) = -1, \forall \beta \in N_i$; and the rest of entries are zero. The number of constraints is the same as the number of atom regions.

The winding number constraints in theory do not depend on the choice of path from each atom region to the outside of the image. However, determining the adjacency relationship numerically can be difficult in the neighborhood of junctions. To minimize the risk of incorrect adjacency estimation, our method chooses a path such that the total number of regions on this path is small, and the edges crossing this path are long. Some examples of paths are shown in Figure 6.

Since region labels are determined by edge labels linearly, the winding number constraints may restrict the set of feasible region labels. However, if region labels are binary, the following proposition shows that the winding number constraints do not restrict solutions of segmentation at all.

Proposition 1. *For any segmentation in which region labels can only be zero or one, there always exists a set of oriented boundaries such that the regional labels equal the winding numbers induced by these boundaries.*

Proof. First of all, we assume that edges do not overlap and each edge is only adjacent to two regions. If this assumption is not valid, the edges can be divided into smaller segments to satisfy the assumption. Then, for an atom region whose label is one, we set a cycle of its adjacent edges in counterclockwise direction to be active. This cycle of edges will induce a

¹If both of the conjugate edges are active, the two regions must share the same label in the same way as when both edges are inactive.

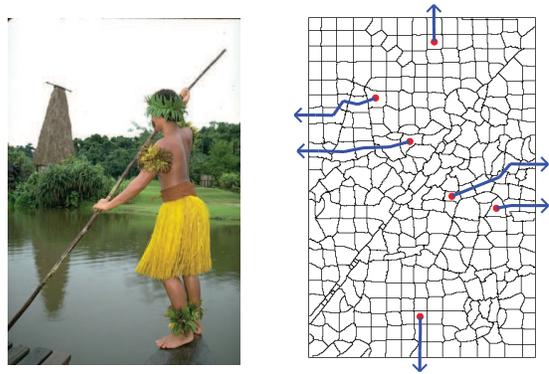


Fig. 6. The left shows an image in BSDS dataset. The right shows paths by which the winding number of the superpixels (red dots) are calculated.

winding number one to this region, and a winding number zero to other regions. Since edges are not shared by more than two regions, this operation can be done to every atom region without conflict. Consequently, every atom region in the foreground has a winding number one. Last, conjugate edges which are both active can be removed without affecting the winding number of any region. Therefore, the final contour map is consistent with the given segmentation. \square

V. APPLICATION TO RATIO-BASED ENERGY FUNCTIONS

Many energy functions have been proposed for image segmentation and contour grouping. Here, we pay special attention to the ratio-based contour detection and segmentation methods which have been studied in [49], [44], [25], [39]. One advantage of ratio-based energy functions is that they have little bias towards either large or small regions. Also, they are not biased towards equal partitions such as normalized cuts. Another advantage is that these energy functions are relatively easy to optimize.

Therefore, ratio-based energy functions are used as an example to demonstrate the effectiveness of our winding number technique. Section V-A shows how contour cues are integrated with region similarity cues through our constraints. Section V-B shows how the curvature cue is integrated. Section V-C extends our method to incorporate user interaction. Implementation details are presented in Section V-E. We also discuss other energy functions in Section V-F.

A. Incorporation of region similarity cue

The contour-based energy function our method adopts is the ratio between contour gaps and areas of foreground, defined as ([44]):

$$\frac{E_B(\mathbf{y})}{A(\mathbf{x})} \quad (11)$$

The boundary term measures the total gap length within contours:

$$E_B(\mathbf{y}) = \alpha_b \sum_i v_i y_i \quad (12)$$

where v_i is the gap length in edge i . The parameter α_b controls strength of the boundary term. This term will favor foreground

objects with salient boundaries. The denominator is the total areas of foreground:

$$A(\mathbf{x}) = \sum_i a_i x_i \quad (13)$$

where a_i is the area of region i . In [44], areas are converted into second edge weights of a graph, and an optimal solution is obtained by solving a graph cycle-finding problem. A problem with Eq (11) is that there may be strong distracting contours inside an object or in background. Here a region similarity term E_R is added to improve robustness to such noise. The new objective function is defined as:

$$E(\mathbf{x}, \mathbf{y}) = \frac{E_R(\mathbf{x}) + E_B(\mathbf{y})}{A(\mathbf{x})} \quad (14)$$

The new addition, the region term is defined as the sum of the affinity measures between figure and background superpixels:

$$E_R(\mathbf{x}) = \alpha_r \sum_{(i,j) \in P_R} w_{ij} |x_i - x_j| \quad (15)$$

where P_R denotes pairs of regions whose distance is smaller than a threshold. The weight w_{ij} encodes the color difference between regions i and j . α_r is a parameter to control the strength of the region term. This term favors large figure-ground contrast.

To ensure region-contour consistency, we use three sets of constraints. The first two sets of constraints are the continuity constraints Eq (5) and the winding number constraints Eq (4). They have been discussed in Section IV. Following the binary label assumption, we limit any region and edge label to be binary. In sum, our ratio-based segmentation model is as follows:

$$\begin{aligned} \min_{\mathbf{x}, \mathbf{y}} \quad & \frac{E_R(\mathbf{x}) + E_B(\mathbf{y})}{A(\mathbf{x})} \\ \text{s.t.} \quad & \Phi_W(\mathbf{x}, \mathbf{y}) = 0 \\ & \Phi_C(\mathbf{y}) = 0 \\ & \mathbf{x} \in \mathcal{X}, \mathbf{y} \in \mathcal{Y} \end{aligned} \quad (16)$$

where the label spaces are defined as $\mathcal{X} = \{0, 1\}^{N_r}$ and $\mathcal{Y} = \{0, 1\}^{N_e}$. Although Eq (16) is good enough for ensuring the region-contour consistency, the formulation can be further simplified by replacing region labels with edge labels using Eq (10), which leads to an energy function of edge variables.

B. Incorporation of curvature cues

Recognized as the Gestalt law of good continuity, the human vision system has a preference for grouping smooth contours together. Our method can be extended to take into account of contour smoothness. The smoothness of contours can be measured by integral of squared curvature. Let P_E denote the indices of all pairs of edges sharing one vertex. The binary junction variable z_{ij} is associated with the junction formed by edge y_i and y_j . Let $\mathbf{z} = \{z_{ij} | (i, j) \in P_E\}$ denote all N_j junction variables, and $\mathcal{Z} = \{0, 1\}^{N_j}$ is the label space of junction variables. In our model, the total curvature cost is defined as:

$$E_C(\mathbf{z}) = \alpha_c \sum_{(i,j) \in P_E} u_{ij} z_{ij} \quad (17)$$

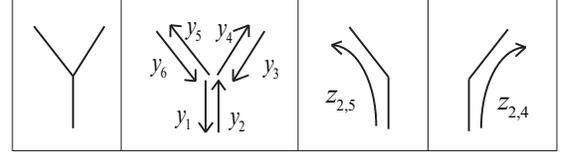


Fig. 7. Our junction model. The first figure shows one junction detected in the image. The second figure shows the 6 variables representing the associated edges. The third and fourth figures show two possible L-junctions if edge y_2 is active.

where the parameter α_c controls the strength of the curvature term as a whole. The curvature weight u_{ij} is the sum of squared curvature along both edges. In [25], only curvature costs within edge fragments are taken into account. However, our curvature term also penalizes sharp turns at junctions.

Our junction model is illustrated in Figure 7. To ensure correct junction configuration, junction constraints are devised. They are denoted as $\Phi_J(\mathbf{y}, \mathbf{z}) \leq 0$. These constraints are adapted from the connectedness constraints for undirected edges [30]. These constraints consist of two parts. First, every active edge should form transition to at least one edge whose tail connects to the head of the current edge. Second, every junction variable can be active only when both of its associated edges are active. The junction constraints are translated into the following linear inequities:

$$\sum_{j | (i,j) \in P_E} z_{ij} \geq y_i, \quad \forall i \quad (18)$$

$$z_{ij} \leq y_i, \quad \forall (i, j) \in P_E \quad (19)$$

$$z_{ij} \leq y_j, \quad \forall (i, j) \in P_E \quad (20)$$

The inequities (18) correspond the first part of junction constraints. The inequities (19) (20) correspond to the second part of junction constraints. In sum, the energy minimization problem is as follows:

$$\begin{aligned} \min_{\mathbf{x}, \mathbf{y}, \mathbf{z}} \quad & \frac{E_R(\mathbf{x}) + E_B(\mathbf{y}) + E_C(\mathbf{z})}{A(\mathbf{x})} \\ \text{s.t.} \quad & \Phi_W(\mathbf{x}, \mathbf{y}) = 0 \\ & \Phi_C(\mathbf{y}) = 0 \\ & \Phi_J(\mathbf{y}, \mathbf{z}) \leq 0 \\ & \mathbf{x} \in \mathcal{X}, \mathbf{y} \in \mathcal{Y}, \mathbf{z} \in \mathcal{Z} \end{aligned} \quad (21)$$

The effectiveness of adding curvature terms is illustrated by the example in Figure 8. The left is an input image. A star shape is favored by the boundary term and the region term due to its stronger contrast to background. Therefore, our model without the curvature term chooses the star as shown in the middle of Figure 8. When the curvature term is added, the smoother disk is extracted as shown in the right of Figure 8.

C. Incorporation of user interaction

Our method described so far only uses low and mid-level grouping cues. It is recognized that the use of high-level information can make a grouping process more robust [5]. A major source of high-level information is user interaction.

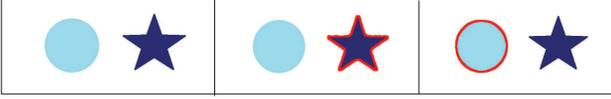


Fig. 8. An example in which the curvature term affects our model's output. The first image is an input image. The second image shows our method's output without the curvature term. Last image shows the output under the influence of the curvature term. (Best viewed in color.)

Our method can be conveniently extended for this purpose. We assume that an user has specified some of edge and region labels by clicking on an input image. Our model can represent these inputs as a set of constraints. For examples, let R_F stand for the regions which the user considers as foreground, and R_B denote those considered as background. The corresponding constraints are:

$$x_i = 1, \quad \forall i \in R_F \quad (22)$$

$$x_i = 0, \quad \forall i \in R_B \quad (23)$$

D. Inference by linear relaxation

The energy minimization problems Eq (16) and Eq (21) are nonlinear integer programs. In the following, they are relaxed into linear programs which can be solved in polynomial time. Since Eq (16) is a special case of Eq (21), we only need to discuss the latter. First, the domain of all labels is relaxed to be interval $[0, 1]$. Second, note that all terms in the objective function are linear except for E_R which is the sum of the absolute values according to Eq (15). Each absolute value $|x_i - x_j|$ is replaced by a variable t_{ij} , and two constraints which are $t_{ij} > x_i - x_j$, and $t_{ij} > x_j - x_i$. Then our model becomes a standard linear fractional program (24). Let the real-valued vector ξ here refer to all the variables, i.e. $\xi^T = [\mathbf{x}^T \ \mathbf{y}^T \ \mathbf{z}^T \ \mathbf{t}^T]$.

Since the denominator representing total areas is strictly positive, the fractional program can be transformed into a linear program [8]. This method is also used in a contour grouping work [20]. In general, the linear fractional program is written as follows:

$$\begin{aligned} \min_{\xi} \quad & \frac{c^T \xi + d}{e^T \xi + f} \\ & A\xi = b \\ & \xi \geq 0 \end{aligned} \quad (24)$$

where A to f are constants. The denominator is positive, i.e. $e^T \xi + f > 0$. Let $\eta = \frac{\xi}{e^T \xi + f}$, $\tau = \frac{1}{e^T \xi + f}$, then the equivalent linear program is:

$$\begin{aligned} \min_{\eta, \tau} \quad & c^T \eta + d\tau \\ & A\eta = b\tau \\ & \eta \geq 0 \\ & e^T \eta + f\tau = 1 \end{aligned} \quad (25)$$

The solution of the fractional program can be obtained as $\xi = \eta/\tau$. In general, ξ is not necessarily integral. However, in our experiments, solutions are usually very close to be integral.

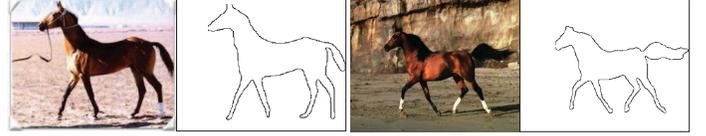


Fig. 9. Weizmann horse dataset. The horse images are in the first and third columns, and the corresponding groundtruth contours are in the second and fourth columns.

E. Implementation details

The boundary gap measure v_i in Eq (12) equals the number of edge pixels in the segment minus the sum of the probabilities of each edge pixel being a true contour point. The probability is estimated according to [25]. The region affinity measure w_{ij} in Eq (15) is the sum of affinity values of all pairs of pixels in these two regions, i.e. $w_{ij} = \sum_{p,q|p \in i, q \in j} w(p, q)$. The pixel-wise weight $w(p, q)$ is computed based on the similarity of colors and locations using a RBF kernel.

We use LP_SOLVE library to solve the linear programming problem. An image is usually oversegmented into about 300 superpixels, and our algorithm consists of ten to twenty thousand variables and several thousands of constraints. LP_SOLVE solves the problem in about twenty seconds on a modest laptop with Intel 2G Centrino 2 core processor/3G RAM.

F. Extension to other objective functions

Other well established objective functions such as normalized cuts can also be transformed into energy functions based on edge labels. Although this objective function is more difficult to optimize, it is included in this paper for discussions. The objective function of normalized cuts [42] is defined as:

$$\min_{\mathbf{x} \in \mathcal{X}} \frac{\sum_{i,j} w_{ij} |x_i - x_j|}{(\sum_i w_i x_i)(\sum_i w_i (1 - x_i))} \quad (26)$$

where w_{ij} are affinity between superpixel i and j . The parameter $w_i = \sum_j w_{ij}$ denotes the volume of x_i . \mathbf{x} denotes all the region labels. Our transformation leads to the following edge-based segmentation problem:

$$\begin{aligned} \min_{\mathbf{y}} \quad & \frac{\sum_{i,j} w_{ij} |m_i^T \mathbf{y} - m_j^T \mathbf{y}|}{(\sum_i w_i m_i^T \mathbf{y})(\sum_i w_i (1 - m_i^T \mathbf{y}))} \\ \text{s.t.} \quad & \Phi_C(\mathbf{y}) = 0 \\ & M\mathbf{y} \in \mathcal{X} \\ & \mathbf{y} \in \mathcal{Y} \end{aligned} \quad (27)$$

VI. EXPERIMENTS

Our method is compared with several related methods on the Weizmann horse dataset [7] and the Weizmann segmentation dataset [1]. Then our method is extended to incorporate user interaction which is effective for more complex BSDS dataset [5].

A. Tests on the Weizmann horse dataset

As shown in Figure 9, this image set contains salient horses in the middle of images. However, obtaining complete contours of horses is still challenging due to several reasons. For example, there are strong distracting contours in background and inside horse regions. True contours on the other hand may be faint or missing because of low contrast. Since horses usually are not camouflaged, we expect the incorporation of region cues to be helpful in obtaining cleaner contours.

Our method is compared with several closely related methods. To demonstrate the effectiveness of cue combination, the region-based normalized cuts method (Ncuts) [42] and the superpixel grouping method (SC) [25] are used for qualitative comparisons. Ncuts is based on pairwise regional affinities, similar to the region term used in our method. SC implements the ratio contour method [44] (RC) in superpixel domain, and our method uses the same superpixel implementation as SC. For quantitative comparisons, we also include RC, and GPAC [45] which is one of the best variational methods for cue combination.

Both SC and our method are initialized on Pb detection results. Ncuts and GPAC implementations are obtained from authors' websites. The parameters of our method are adjusted on a validation set. The results are shown in Figure 10. SC outputs ten solutions for each image, the one with the highest F-value is shown in the figure. We can see that our outputs better separate the horse regions from background. In SC's outputs, horse legs are often connected as a single blob region. However, region similarity cues used in our model help distinguish background regions from foreground regions. Two-way Ncuts often cuts out a homogeneous background area. Ten-way Ncuts, however, tends to produce spurious edges (e.g. those in the sky and grass). These results show that either contour cues or region cues alone are not enough for segmenting salient foreground regions. Note that our results in Figure 10 appear to be a single contour due to our choice of the objective function. The winding number constraints, however, do not require the solution to be a Jordan curve.

We quantitatively evaluate related methods using F-measure on 100 horse images. The F-value of each solution is computed by comparing its segmentation mask with the groundtruth mask. The F-value is close to 100% if and only if the detected contours are close to the groundtruth contours, and the F-value declines gradually when the output contours deviate from groundtruth. For methods producing more than one solutions per image (e.g. SC, Ncuts), the overall F-value of an image is the best F-value of all the solutions. Figure 11 shows the F-values of SC reported in [25] which converges to 79.7% with 10 solutions per image. In our experiment, it converges to 76.5%. However, in either [25] or our experiment, the performance is much worse when the number of solution is small. Our model outputs only one solution for each image and achieves an F-value of 74.1%. RC reaches an F-value of 68.1%. We let Ncuts to produce two or ten solutions for one image by partition the image into two or ten segments. The F-value of two-way Ncuts is 43.7% while that of ten-way Ncuts is 64.9%. The F-value of GPAC is 54.5%. Using the bound-

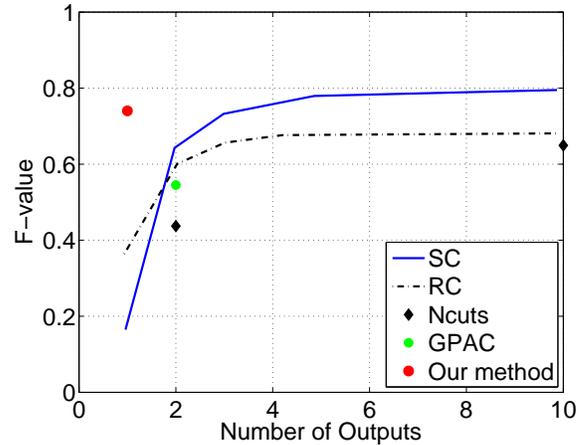


Fig. 11. The F-values of related methods on Weizmann horse dataset. SC method achieves high F-value with 10 the solutions. However, our method is better when considering only one solution. (Best viewed in color.)

ary term alone, our method achieves an F-value of 64.8%, demonstrating that the region homogeneity cue is effective for this dataset. Note that with additional constraints [11], our method could output more solutions at the expense of increased computation time.

B. Tests on the Weizmann segmentation dataset

Our method is also tested on the Weizmann segmentation dataset [1] which contains one hundred single-object images and one hundred images with two objects. Our method is adjusted on the two-object images and is tested on the single-object images. We follow the same protocol as in the previous section. Qualitatively, our method can produce good results as shown in Figure 12. SC, Ncuts, RC, and GPAC methods are used in the comparison by F-measure. When compared with the best of ten solutions, our method's single solution F-value (78.02%) is inferior to SC (87.19%), and is comparable to RC (77.82%). However, as shown in Figure 13, our method has clear advantage for the single output. In addition, for some images our method can produce better outputs than the best solutions of SC and Ncuts. The F-value of two-way Ncuts is 45.7%, and that of ten-way Ncuts is 56.8%. The F-value of GPAC is 63.6%. We have analyzed the contributions of each term. When there is only boundary term in the objective function the F-value of our solutions reduce to 76%. Our method using the boundary term and the region term but not the curvature term achieves 78%. The improvement by the curvature term is largely in details which do not affect F-value much.

C. Incorporating user interaction on BSDS300 dataset

To demonstrate the effectiveness of incorporating user interaction, our baseline method is firstly evaluated on the BSDS300 dataset which contains a variety of urban and natural scene images. The BSDS dataset is more challenging because many images contain multiple foreground objects and background clutter. Some sample results are shown in



Fig. 10. Comparison with the superpixel closure method (SC) and the normalized cuts method (Ncuts). The first column shows the input images. The second column shows our results. The third column shows the SC results. Only the best solution (with the highest F-value) of each image is shown out of 10 solutions. The fourth column shows the 2-way segmentation results by Ncuts. The fifth column shows 10-way segmentation results by Ncuts. The last column shows GPAC results. (Best viewed in color.)

Figure 14, and they match our perception of salient objects. As shown in the first row of Figure 15, the addition of region cues sometimes bias results towards homogenous parts of objects. When several objects have similar colors, our solution often cannot differentiate these objects, as shown in the second row. In addition, we observe the difficulty of detecting camouflaged objects for which neither cue is effective. The satisfactory results are only produced for images in which the foreground object/region is evident. Overall, these observations confirm that our results indeed reflect both region and contour cues

but high-level cues are needed for more accurate results.

Our method is extended to use interaction as follows. The automatic contour detection result of an input image is first shown to a user. An interface is then provided to record the user's clicks on the input image. Mouse positions of the left clicks are assumed to be in the figural region, and those of right clicks to be in the background. We detect the atom regions which are clicked on. The labels of these regions are constrained to be zero or one depending on the type of clicks. These constraints derived from user interaction are added into

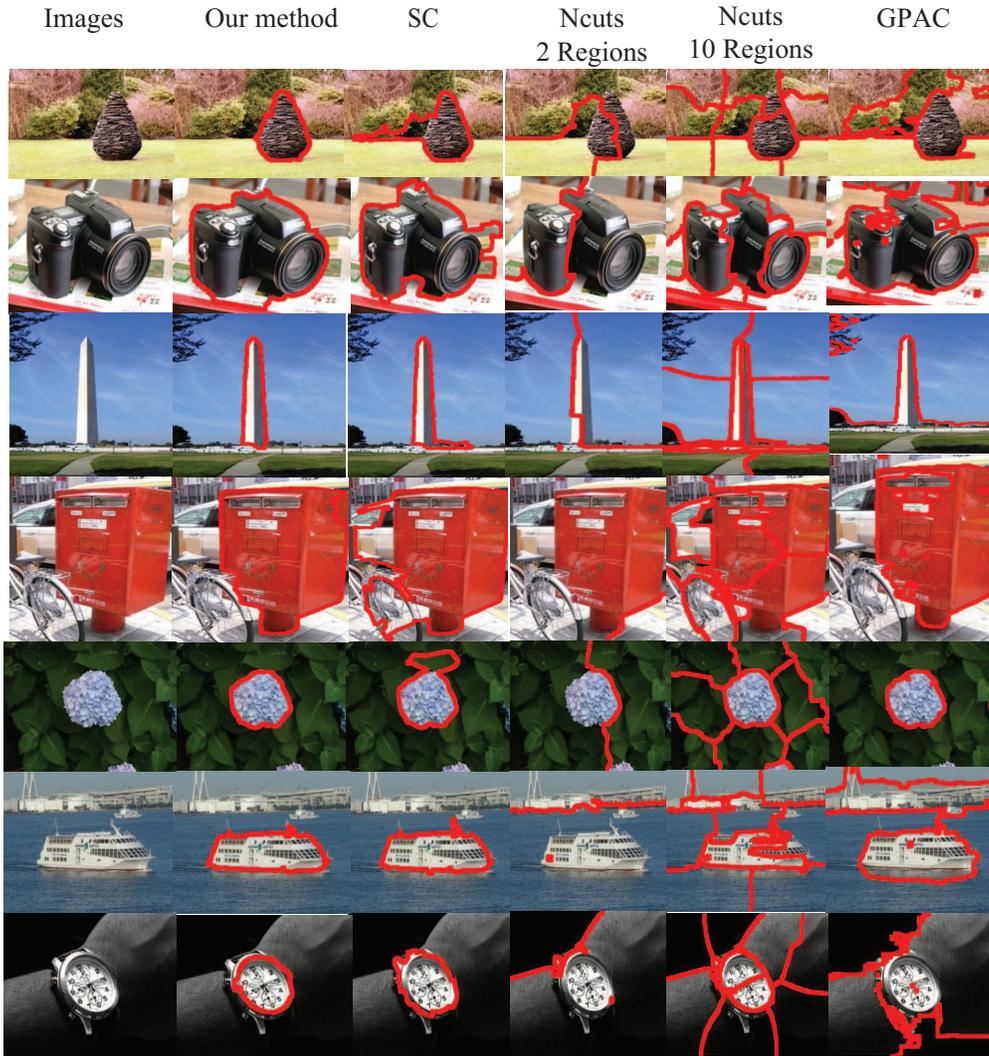


Fig. 13. Some images in Weizmann segmentation dataset for which our method’s contour outputs are either more concentrated on the objects (row 1 to row 5) or smoother (last two rows), due to the influence of the region and curvature terms. The first column shows the input images. The second column shows our results. The SC results displayed in the second column are the best ones (in terms of F-value) out of ten solutions. The Ncuts results are shown in fourth and fifth columns. GPAC results are in the last column.

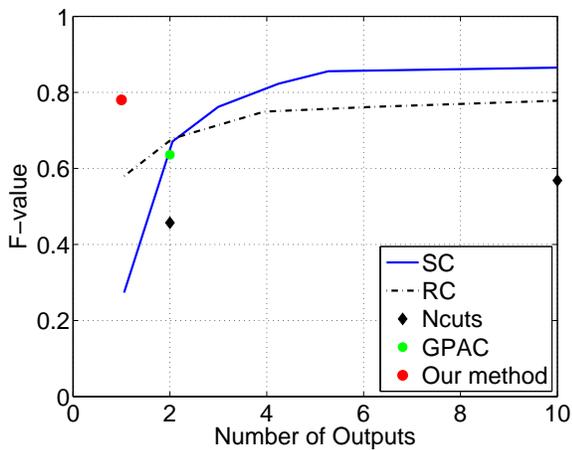


Fig. 12. The F-values of related methods on Weizmann segmentation dataset (single-object images).

the system and a new solution is obtained and presented to the user. The user can provide additional inputs until the solution becomes satisfactory. Figure 16 shows that our method can find solutions which are consistent with human inputs. Clear improvements are observed on these challenging images due to the high-level inputs.

Results by the popular interactive segmentation methods Grabcuts [38] and Random Walker [15] are also shown for comparison, using codes implemented by Blumenthal², and [4], respectively. The quality of a segmentation result is measured by an index which is the intersection of the segmentation mask and its groundtruth mask over the union of two masks. The value of this index is between zero and one. For these four images, the average indices of our method, Grabcuts, and Random Walker are 0.89, 0.87 and 0.76, respectively. Grabcuts has comparable performance to our method. However, labeling a bounding polygon is more

²<http://grabcut.weebly.com/index.html>

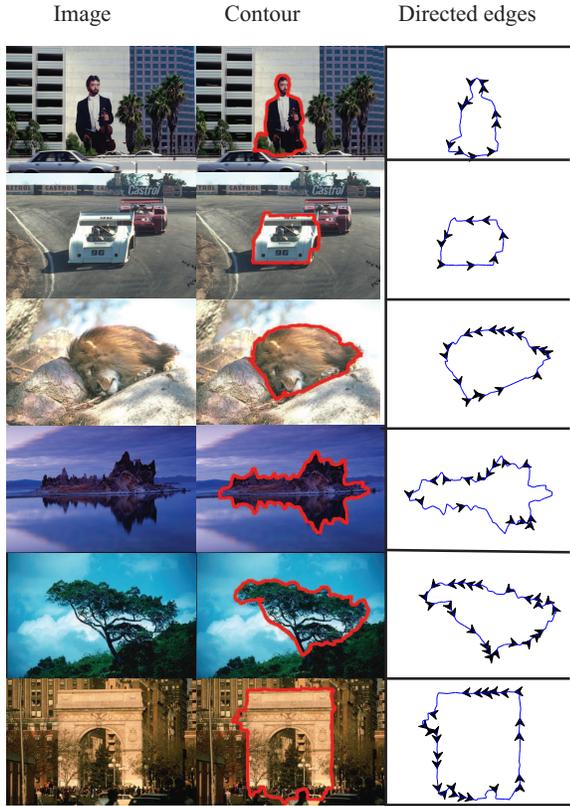


Fig. 14. Sample results on BSDS300 dataset. The first column is the input images. The second column is the output contour overlaid on the input images. The third column shows the directed active edges. (Best viewed in color.)



Fig. 15. Examples of segmentation bias resulted from the homogeneity cue. For images in the first row, the extracted contour focused on homogenous regions rather than whole objects. For images in the second row, our method cannot separate different objects with similar color. (Best viewed in color.)

laborious than labeling several seeds. Random Walker method also uses seeds as input. However, it requires a larger set of seeds in order to obtain good results for those images. In this experiment, the seed set for Random Walker is a superset of ours. Besides human input, our method can be extended to use object detector’s output.

1) *Seeds selection*: Our method’s results depend on the user input, as shown in Figure 17. However, all these results show improvements over the one without interaction. This demonstrates the robustness of our method under input variations. An important question is how to select an optimal set of seeds. Our experience is that the seeds are better placed

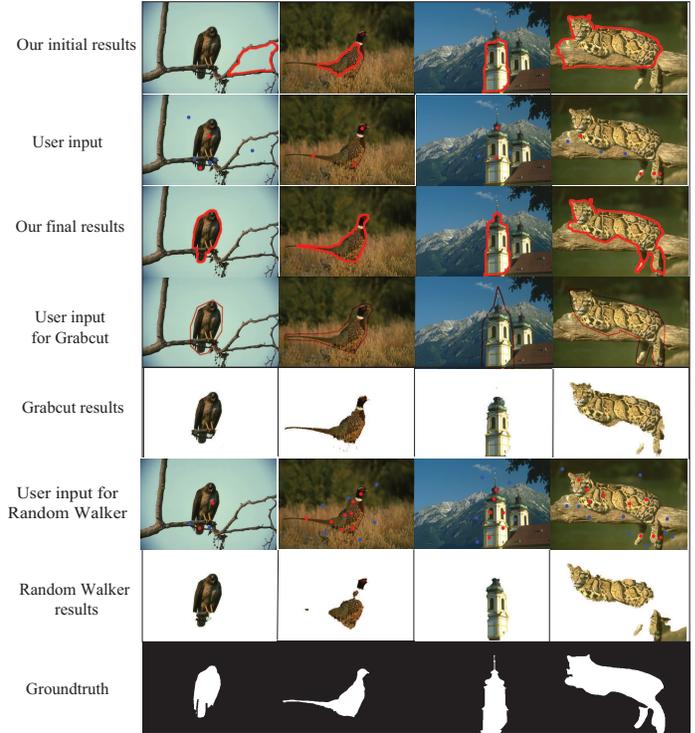


Fig. 16. Improved contour extraction results by user interaction. The first row shows our outputs not using interaction. The second row shows user inputs. The red dots denote the regions in foreground and the blue dots denote the regions in the background. The third row shows our outputs under the guidance of the user inputs. The user input for GrabCut, and its results are shown in the fourth and fifth rows, respectively. The next two rows show input for Random Walker method and its results. The last row shows the groundtruth segmentation masks. (Best viewed in color.)

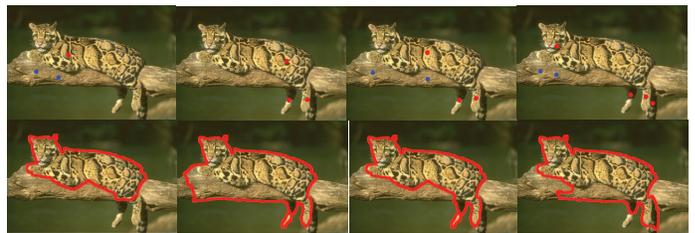


Fig. 17. Different user inputs (first row), and the corresponding segmentation results (second row) by our method.

near the boundaries where the segmentation is not accurate. To save human effort, our interactive segmentation method can feedback the current segmentation result to the user after every click. In this way, the user can correct the segmentation errors greedily. In contrast, an user must provide a whole set of polygon vertices to the Grabcuts before getting any feedback.

VII. CONCLUSION AND FUTURE WORK

This paper introduces a method for combining region and contour representations efficiently, based on the winding number concept. This model is simple and appealing, as it only involves a compact set of linear constraints to ensure

the consistency of both representations. As an application of this method, region similarity cues and region-based user interaction are added into our ratio-based contour detection framework, and lead to improved results. This method is further extended to incorporate user interaction as a special kind of region cues. In future, we believe that more sophisticated design of region/contour cues could help to extract contours of complicated objects. We are also interested in finding efficient optimization methods for other objective functions and extend this method to the multiple-label case.

ACKNOWLEDGMENT

This work is in part funded by Australian Research Council through ARC Discovery grants. NICTA is funded by Australian Government and by ARC. We thank Srimal Jayawardena for proofreading.

REFERENCES

- [1] S. Alpert, M. Galun, R. Basri, and A. Brandt, "Image segmentation by probabilistic bottom-up aggregation and cue integration," in *Proc. CVPR*, 2007.
- [2] S. Alpert, M. Galun, A. Brandt, and R. Basri, "Image segmentation by probabilistic bottom-up aggregation and cue integration," *IEEE TPAMI*, vol. 34, no. 2, pp. 315–327, 2012.
- [3] B. Andres, J. H. Kappes, T. Beier, U. Köthe, and F. Hamprecht, "Probabilistic image segmentation with closedness constraints," in *Proc. ICCV*, 2011.
- [4] S. Andrews, G. Hamarneh, and A. Saad, "Fast random walker with priors using precomputation for interactive medical image segmentation," in *Lecture Notes in Computer Science, Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2010, pp. III:9–16.
- [5] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE TPAMI*.
- [6] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.
- [7] E. Borenstein and S. Ullman, "Class-specific, top-down segmentation," in *Proc. ECCV*, 2002, pp. 109–124.
- [8] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.
- [9] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE TPAMI*, 2001.
- [10] J. Canny, "A computational approach to edge-detection," *IEEE TPAMI*, vol. 8, no. 6, pp. 679–698, 1986.
- [11] J. Carreira and C. Sminchisescu, "Cpmc: Automatic object segmentation using constrained parametric min-cuts," *IEEE TPAMI*, vol. 34, no. 7, pp. 1312–1328, 2012.
- [12] J. H. Elder and S. W. Zucker, "Computing contour closure," in *Proc. ECCV*, 1996, pp. 399–412.
- [13] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *IJCV*, pp. 167–181, 2004.
- [14] J. Gallier and D. Xu, "The fundamental group, orientability," in *A Guide to the Classification Theorem for Compact Surfaces*, ser. Geometry and Computing. Springer, 2013, vol. 9.
- [15] L. Grady, "Random walks for image segmentation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 11, pp. 1768–1783, 2006.
- [16] B. Hariharan, P. Arbelaez, L. Bourdev, S. Maji, and J. Malik, "Semantic contours from inverse detectors," in *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011, pp. 991–998.
- [17] X. He, R. S. Zemel, and M. Á. Carreira-Perpiñán, "Multiscale conditional random fields for image labeling," in *Proc. CVPR*, 2004, pp. 695–702.
- [18] A. Jacobson, L. Kavan, , and O. Sorkine-Hornung, "Robust inside-outside segmentation using generalized winding numbers," *ACM SIGGRAPH*, vol. 32, no. 4, pp. 33:1–33:12, 2013.
- [19] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes - active contour models," *IJCV*, vol. 1, no. 4, pp. 321–331, 1987.
- [20] I. Kokkinos, "Highly accurate boundary detection and grouping," in *Proc. CVPR*, 2010, pp. 2520–2527.
- [21] —, "Boundary detection using f-measure-, filter- and feature- (f3) boost," in *Proc. ECCV*, 2010, pp. 650–663.
- [22] I. Kovacs and B. Julesz, "A closed curve is much more than an incomplete one - effect of closure in figure ground segmentation," *PNAS, USA*, vol. 90, no. 16, pp. 7495–7497.
- [23] T. Leung and J. Malik, "Contour continuity in region based image segmentation," in *Proc. ECCV*, 1998, pp. 544–559.
- [24] A. Levinshtein, A. Stere, K. Kutulakos, D. Fleet, S. Dickinson, and K. Siddiqi, "Turbopixels: Fast superpixels using geometric flows," *IEEE TPAMI*, vol. 31, no. 12, pp. 2290–2297, 2009.
- [25] A. Levinshtein, C. Sminchisescu, and S. Dickinson, "Optimal contour closure by superpixel grouping," in *Proc. ECCV*, 2010, pp. 480–493.
- [26] S. Mahamud, L. R. Williams, K. K. Thornber, and K. L. Xu, "Segmentation of multiple salient closed contours from real images," *IEEE TPAMI*, vol. 25, no. 4, pp. 433–444, 2003.
- [27] D. R. Martin, C. C. Fowlkes, and J. Malik, "Learning to detect natural image boundaries using local brightness, color, and texture cues," *IEEE TPAMI*, vol. 26, no. 5, pp. 530–549, 2004.
- [28] M. McIntyre and G. Cairns, "A new formula for winding number," *Geometriae Dedicata*, vol. 46, no. 2, pp. 149–159, 1993.
- [29] A. L. F. Meister, *Generalia de genesi figurarum planarum et inde pendentibus earum affectionibus*, 1769.
- [30] Y. Ming, H. Li, and X. He, "Connected contours: A new contour completion model that respects the closure effect," in *Proc. CVPR*, June 2012, pp. 829–836.
- [31] —, "Winding number for region-boundary consistent salient contour extraction," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*. IEEE, 2013, pp. 2818–2825.
- [32] T. Needham, *Visual Complex Analysis*. Oxford University Press, USA, Feb. 1999.
- [33] F. Nicolls and P. H. Torr, "Discrete minimum ratio curves and surfaces," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 2133–2140.
- [34] S. E. Palmer, J. L. Brooks, and R. Nelson, "When does grouping happen?" *Acta Psychologica*, vol. 114, no. 3, pp. 311–330, 2003.
- [35] G. Papari and N. Petkov, "Edge and line oriented contour detection: State of the art," *Image and Vision Computing*, vol. 29, no. 2-3, pp. 79–103, 2011.
- [36] P. Perona and J. Malik, "Detecting and localizing edges composed of steps, peaks and roofs," *Third ICCV*, pp. 52–57, 1990.
- [37] S. B. Piotr Dollar, Zhuowen Tu, "Supervised learning of edges and object boundaries," in *Proc. CVPR*, 2006, pp. 1964–1971.
- [38] C. Rother, V. Kolmogorov, and A. Blake, "'grabcut': interactive foreground extraction using iterated graph cuts," in *ACM SIGGRAPH*, 2004, pp. 309–314.
- [39] T. Schoenemann, S. Masnou, and D. Cremers, "The elastic ratio: Introducing curvature into ratio-based image segmentation," *IEEE TIP*, pp. 2565–2581, 2011.
- [40] T. Schoenemann, F. Kahl, S. Masnou, and D. Cremers, "A linear framework for region-based image segmentation and inpainting involving curvature penalization," *IJCV*, vol. 99, no. 1, pp. 53–68, 2012.
- [41] J. A. Sethian, *Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science*. Cambridge university press, 1999, vol. 3.
- [42] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE TPAMI*, vol. 22, pp. 888–905, 2000.
- [43] J. S. Stahl and S. Wang, "Globally optimal grouping for symmetric closed boundaries by combining boundary and region information," *IEEE TPAMI*, vol. 30, no. 3, pp. 395–411, 2008.
- [44] —, "Edge grouping combining boundary and region information," *IEEE TIP*, vol. 16, no. 10, pp. 2590–2606, 2007.
- [45] B. Sumengen and B. S. Manjunath, "Graph partitioning active contours (gpac) for image segmentation," *IEEE TPAMI*, pp. 509–521, 2004.
- [46] M. Tabb and N. Ahuja, "Multiscale image segmentation by integrated edge and region detection," *IEEE TIP*, vol. 6, no. 5, pp. 642–655, May 1997.
- [47] L. Vincent and P. Soille, "Watersheds in digital spaces: an efficient algorithm based on immersion simulations," *IEEE transactions on pattern analysis and machine intelligence*, vol. 13, no. 6, pp. 583–598, 1991.
- [48] D. B. Walther, B. Chai, E. Caddigan, D. M. Beck, and L. Fei-Fei, "Simple line drawings suffice for functional mri decoding of natural scene categories," *Proceedings of the National Academy of Sciences*, vol. 108, no. 23, pp. 9661–9666, 2011.
- [49] S. Wang, T. Kubota, J. M. Siskind, and J. Wang, "Salient closed boundary extraction with ratio contour," *IEEE TPAMI*, vol. 27, no. 4, pp. 546–561, 2005.

- [50] S. Wang and J. M. Siskind, "Image segmentation with ratio cut," *IEEE TPAMI*, vol. 25, pp. 675–690, 2003.
- [51] H. Whitney, "On regular closed curves in the plane," *Compositio Mathematica*, vol. 4, pp. 276–284, 1937.
- [52] L. R. Williams and K. K. Thornber, "A comparison of measures for detecting natural shapes in cluttered backgrounds," *IJCV*, vol. 34, no. 2-3, pp. 81–96, 1999.
- [53] X. Xie and M. Mirmehdi, "Rags: Region-aided geometric snake," *IEEE TIP*, pp. 640–652, 2004.
- [54] S. X. Yu, T. S. Lee, and T. Kanade, "A hierarchical markov random field model for figure-ground segregation," *EMMCVPR*, 2001.
- [55] L. Zhang and Q. Ji, "Image segmentation with a unified graphical model," *IEEE TPAMI*, vol. 32, no. 8, pp. 1406–1425, 2010.
- [56] S. Zheng, A. Yuille, and Z. Tu, "Detecting object boundaries using low-, mid-, and high-level information," *Computer Vision and Image Understanding*, vol. 114, no. 10, pp. 1055–1067, 2010.
- [57] C. Zitnick and D. Parikh, "The role of image understanding in contour detection," in *Proc. CVPR*, june 2012, pp. 622–629.