

Winding Number for Region-Boundary Consistent Salient Contour Extraction

Yansheng Ming
ANU
Canberra, ACT, Australia
yansheng.ming@anu.edu.au

Hongdong Li
ANU and NICTA
Canberra, ACT, Australia
hongdong.li@anu.edu.au

Xuming He
NICTA and ANU
Canberra, ACT, Australia
xuming.he@nicta.com.au

Abstract

This paper aims to extract salient closed contours from an image. For this vision task, both region segmentation cues (e.g. color/texture homogeneity) and boundary detection cues (e.g. local contrast, edge continuity and contour closure) play important and complementary roles. In this paper we show how to combine both cues in a unified framework. The main focus is given to how to maintain the consistency (compatibility) between the region cues and the boundary cues. To this ends, we introduce the use of winding number—a well-known concept in topology—as a powerful mathematical device. By this device, the region-boundary consistency is represented as a set of simple linear relationships. Our method is applied to the figure-ground segmentation problem. The experiments show clearly improved results.

1. Introduction

Salient contours in an image play a central and important role in object perception and scene understanding. Often these salient contours correspond to semantically meaningful contents in the image, such as object boundaries. Therefore, extracting a few clean and meaningful contours may simplify subsequent high-level image understanding tasks. However, salient contour extraction is a challenging task as it involves both region and boundary information, requiring integration of bottom-up image cues and top-down semantic priors. In particular, many perceptual grouping laws such as proximity, continuity, closure and symmetry, *etc.* have been shown to be critical in such process.

Contour extraction has been approached through two complementary directions: one is to treat the problem as a (2D) region segmentation task [18, 23, 7], and the other focuses on the intrinsic 1D contour detection aspect of the problem (such as snake/level-set methods). There has been a trend of jointly using, or combining both contour cues and region cues. In the continuous domain, active contour model was adapted to use both region and contour

cues [20] [25]. In the discrete domain, contour cues (such as curvature) have been introduced to region-segmentation methods (e.g., [19], *intervening contour* approach in [10], SC [11]). However, many of these methods either lack explicit region-contour interaction, or rely on heuristic or an unduly complicated model [27].

This paper aims to develop a more consistent approach to salient contour extraction that tightly integrates both region cues and boundary cues. Our insight is that, to achieve jointly utilizing both aspects of image cues, using a simple linear combination of a region objective function and a contour objective function is not sufficient. A key issue, that must be taken into account, is the conditions under which the consistency (or compatibility) between the region variables and the edge variables is satisfied. Otherwise, meaningless solutions (such as an object with fragmented boundary) may occur. However, it is recognized that, enforcing such consistency is not a trivial task. Paper [2] argued that in order to ensure the closedness condition, exponentially many constraints are needed. A recent work [16] gave a linear programming framework for enforcing the boundary-region consistency, again, relying on a large number of linear inequality constraints.

In this work, we propose a novel and simpler method to describe the region-contour consistency relationship, borrowing “winding number” as a handy concept from the mathematical field of topology. Our key intuition that motivates this model is a well-known fact: contour and region form certain kind of “duality” relationship. 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.

The above observation motivates our winding number approach of this paper. By definition, winding number, which involves a set of closed (but not necessarily simple) planar curves and a point in the plane, refers to the number

of times the curves revolves around this point. The key idea of this paper is that if the region labels are restricted to be the winding numbers of the curves, then the region-contour consistency condition can be effectively captured by a very small set of linear constraints.

Using the winding number concept, we present a new salient contour detection model, integrating the region segmentation cue into the ratio-based contour detection framework [22][19]. In particular, we focus on the foreground segmentation task, in which the winding number can be transformed to region label exactly. Our objective function includes constraints from contour saliency, region similarity and contour smoothness, and can be efficiently solve by linear programming approximately.

Our method is evaluated on the Weizmann horse images and Berkeley segmentation dataset, showing advantages over pure contour or region based approaches. Although this paper focuses on middle-level perceptual grouping, we believe that our method is applicable too for higher level tasks such as object detection where contour cues and region cues are both helpful.

The paper is organized as follows. Section 2 discusses relevant methods in the literature. Section 3 presents the winding number method in detail. In Section 4, our method is used for integrating region cue and contour cues to achieve salient contour extraction. Experimental results are shown in Section 6, and the conclusion is in Section 7.

2. Related work

Our work is closely related to contour grouping and image segmentation. Most of the contour grouping methods start with local edge detection e.g. [13]. Then various mid-level to high-level cues are used to improve the accuracy. For example, contour smoothness prior is popularly employed as a prominent contour cue [12] [24]. Contour closure prior has been addressed in [2] [14].

We will demonstrate our winding number method under the graph partition framework. The general idea of this framework is to model image pixels as a set of coupled graph nodes. Image segmentation problem is transformed into a graph partition problem. Notable methods under this framework include the max-flow [7], normalized cuts [18] and the ratio cut [23]. Moreover, spectral K-means [4], Felzenszwalb *et al.*'s graph-based method [9], SWA method [1] also fall under a general graph partition framework.

There are several previous works which attempted to incorporate both region cues and contour cues. Intervening-contour [10] is one of the early efforts to use local contour strength for region segmentation. This method is built into the state of the art segmentation algorithm [3]. Yu *et al.* [26] also incorporated edge information in a framework based on Markov random field. Tabb and Ahuja integrated both cues

for low level structure detection [21]. GPAC method [20] has shown the flexibility to accommodate both region and contour cues in one energy function. However gradient descend inference method is susceptible to local minimal. Stahl and Wang [19] modified the ratio contour method [22] by replacing the total length with total area as the denominator which resulted in segmentation of more regular shape. Recently, [11] applied the method of [19] to superpixel grouping problem. In setting of interactive segmentation, [16] used a set of local consistency constraints to achieve the boundary region consistency. In contrast, the winding number constraints used by our method is a set of global constraints. The winding number concept not only leads to smaller number of constraints, but also makes the framework potentially applicable to multiple-label segmentation, although not demonstrated in this paper.

Finally, the concept of winding number (rotation index) has been used for ensuring contour topology in [8]. Different from their method, our method use this concept to ensure the region/contour consistency.

3. Winding number representation

This section first presents our salient contour extraction problem setting, which is based on superpixel over-segmentation. Section 3.2 gives a high level description of our method. Section 3.3 presents the main idea of our winding number-based method.

3.1. Basic edge and region hypotheses

We formulate the salient contour extraction problem as an energy minimization problem defined on both region and edge hypotheses. We choose superpixel over-segmentation as a means to provide sufficient edge and region hypotheses. Each superpixel provides an atom region hypothesis. We fit the boundary of each superpixel 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 graph as well.

3.2. High-level description of our method

let $\mathbf{x} = \{x_i | i = 1 \dots N_r\}$ denote the labels of N_r atom regions, and $\mathbf{y} = \{y_j | j = 1 \dots N_e\}$ denote the labels of N_e edges. Each region can have one of the integer labels from a predefined label set. The label space of all region variables is denoted as \mathcal{X} . The edge label space is denoted as $\mathcal{Y} = \{0, 1\}^{N_e}$. Generally, the salient contour detection is formulated as the following energy-minimization problem:

$$\begin{aligned}
\min_{\mathbf{x}, \mathbf{y}} \quad & E(\mathbf{x}, \mathbf{y}) & (1) \\
s.t. \quad & \Phi_W(\mathbf{x}, \mathbf{y}) = 0 & (2) \\
& \Phi_C(\mathbf{x}, \mathbf{y}) = 0 & (3) \\
& \mathbf{x} \in \mathcal{X}, \mathbf{y} \in \mathcal{Y} & (4)
\end{aligned}$$

where the energy function E captures various region priors and contour priors. Particular form of the energy function is not the concern of this work.

Our main contribution of this paper is the introduction and construction of a compact set of “winding number constraints”, Eq (2) in the above formulation, which captures the consistency relationship between edge variables and region variables. The edge continuity constraints Eq (3) is necessary for ensuring the edges forms cycles in the graph. These constraints together ensure that the obtained edge and region labeling result to be topologically valid and semantically meaningful. The specific consistency condition used in this paper requires 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 a part of a closed region boundary, and every region is enclosed by a boundary (or contour). Conversely, violating this condition will lead to the break of contour connectedness/closure condition, as shown in work [2]. Next section will show that this condition is guaranteed by our constraint sets based on the winding number concept.

3.3. Winding number and its fast computation

We realize that the winding number concept, from topological study, provides an elegant and effective means to parameterize the region-contour consistency constraint in image segmentation.

The winding number of a point induced by a closed curve is defined as the number of times this curve travels around the point counterclockwise [15]. For a set of contour, the induced winding number can be defined as the sum of winding numbers induced by every contour. Provable by the celebrated Residue Theorem [15], the winding number of all the image points inside an atom region must be all equal. Based on this remarkable result, we reach our *winding number constraint*, viz.

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

Winding numbers in region segmentation are not unique. Different partitions lead to different winding numbers. Figure-2 illustrates how different contour labels result in different segmentations. We emphasize that the winding

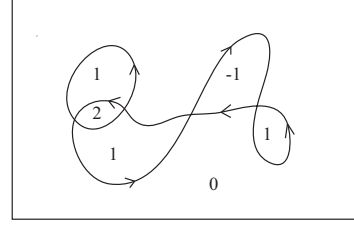


Figure 1. Winding numbers induced by a set of closed contours.

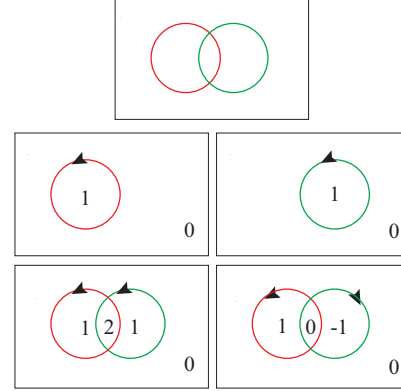


Figure 2. The top row shows an image of two circles. The rest of the figure shows different segmentations induced by different contour orientation. The winding numbers are shown in respective regions. Regions with the same winding number are considered as in the same region.

number constraints, unique or not, guarantee the consistency between region labels and the corresponding contour labels. First of all, the winding numbers of adjacent regions will be different if one of the conjugate edges between them is active.¹ Secondly, the Residue Theorem 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. In conclusion, this winding number scheme does encode the region-contour consistency condition compactly and efficiently.

The benefit of such winding number scheme also lies in that: it leads to a set of linear constraints. This can be made evident by examining the fast computation procedure of winding number computation (c.f. [15]). Given an image, consider everything outside the image frame (image border) is void. We assign a label of zero (0) to that part. Then we draw an arbitrary path starting from inside a region to outside of the image frame, then the winding number of the region equals to the number of edges crossing the path from the right side minus the number of edges crossing from the left side. This fast computation procedure is illustrated

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.

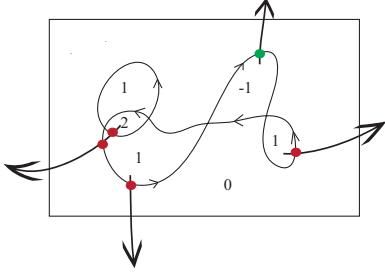


Figure 3. Fast winding number computation. Draw an arbitrary path to outside of the image frame, the winding number of a point equals the number of edges crossing from right (red dot) minus the number of edges crossing from the left (green dot).

in Figure 3. Formally, the winding number of the region i is computed as:

$$x_i = \sum_{\alpha \in P_i} y_\alpha - \sum_{\beta \in N_i} y_\beta, \quad \forall i \quad (5)$$

where P_i and N_i are the edges crossing from right to left, and edges crossing from left to right, respectively. Eq (5) for all atom regions together can be represented as the following winding number constraint, denoted as Φ_W in Eq (2):

$$\mathbf{x} = M\mathbf{y}, \quad (6)$$

where M is a matrix whose entries are 0, 1, or -1 .

To ensure that the extracted contour form cycles, we introduce the edge-continuity constraint:

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

where j is one index of the vertex index set V . j_{in} and j_{out} denote the edges indices heading into and moving out of the vertex j , respectively. These edge continuity constraints for all vertices are denoted as $\Phi_C(\mathbf{y}) = 0$. These constraints say that the net flow at every vertex is zero. For a network without source and sink, it can be shown that all the flows can be decomposed into a set of cycles and the winding number can be computed.²

For a general K-way cut problem, the winding number constraint may restrict the feasible set of region labels. However, for the figure-ground segmentation problem, which is of interest to this paper, the following proposition shows that the winding number constraints do not restrict the solution of segmentation at all.

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

²The degenerate cycle formed by conjugate edges cannot be excluded by these constraints. However, they do not affect correctness of the winding number computation.

Proof. First of all, we assume that edges do not overlap and each edge is only adjacent to two regions. If the 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 counter-clockwise direction to be active. This cycle of edges will induce a 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, the conjugate edges which are both active can be removed without affecting the winding number of any region. Therefore, the resulted contour is the one consistent with the given segmentation. \square

4. Ratio-Contour with region cue

The rest of the paper is focused on solving the figure-ground segmentation problem as a special case of Eq (1). Ratio-based contour detection and segmentation methods have been studied in [22] [19] [11] [17]. Using the contour cue, the objective function can be the ratio of contour gap over total contour length or figural areas. Here, we use ratio-based method as an example to demonstrate the effectiveness of our winding number scheme. In Section 4.1, we will show how the contour gap information is integrated with the region similarity cue. Section 4.2 will explain how curvature cue is integrated.

4.1. Incorporation of region similarity cue

The contour-based energy function our method adopts is a ratio between the contour gap and the areas of foreground, defined as ([19]):

$$\frac{E_B(\mathbf{y})}{A(\mathbf{x})} \quad (8)$$

The boundary term measures the gap in the contour:

$$E_B(\mathbf{y}) = \alpha_b \sum_i v_i y_i \quad (9)$$

where v_i is the gap length in edge i . The parameter α_b controls the strength of the boundary term. This term will favor the foreground with a salient boundary. The denominator is the total areas of the foreground:

$$A(\mathbf{x}) = \sum_i a_i x_i \quad (10)$$

where a_i is the areas of region i . In their work, the areas are converted into second edge weights of a graph, and the optimal solution is obtained by solving a graph cycle-finding problem. The problem with objective function Eq (8) is that there may be strong distracting contours inside the object or in the background. Here region similarity term is added to

increase the accuracy of segmentation. The new objective function is defined as:

$$E(\mathbf{x}, \mathbf{y}) = \frac{E_R(\mathbf{x}) + E_B(\mathbf{y})}{A(\mathbf{x})} \quad (11)$$

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

$$E_R(\mathbf{x}) = \alpha_r \sum_{(i,j) \in P_R} w_{ij} |x_i - x_j| \quad (12)$$

where P_R denotes the set of pairs of region whose distance is smaller than a threshold. The weight w_{ij} of the cut of region i and j encodes the color difference of the two regions. α_r is a parameter to control the strength of the region term as a whole. 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 (3) and winding number constraints Eq (2). They have been discussed in Section 3. For the figure-ground segmentation problem, it is necessary to limit any region and edge label to be zero and one. In sum, our ratio-based segmentation model is as follows:

$$\begin{aligned} \min_{\mathbf{x}, \mathbf{y}} \quad & \frac{E_R(\mathbf{x}) + E_B(\mathbf{y})}{A(\mathbf{x})} \quad (13) \\ \text{s.t.} \quad & \Phi_W(\mathbf{x}, \mathbf{y}) = 0 \\ & \Phi_C(\mathbf{y}) = 0 \\ & \mathbf{x} \in \mathcal{X}, \mathbf{y} \in \mathcal{Y} \end{aligned}$$

where the label spaces are defined as $\mathcal{X} = \{0, 1\}^{N_r}$ and $\mathcal{Y} = \{0, 1\}^{N_e}$. Although Eq (13) is good enough for ensuring the region-contour consistency, the formulation can be further simplified by replacing region labels with edge labels using Eq (6). As a result, our problem formulation depends on edge variables only.

$$\begin{aligned} \min_{\mathbf{y}} \quad & \frac{E_R(M\mathbf{y}) + E_B(\mathbf{y})}{A(M\mathbf{y})} \quad (14) \\ \text{s.t.} \quad & \Phi_C(\mathbf{y}) = 0 \\ & M\mathbf{y} \in \mathcal{X} \\ & \mathbf{y} \in \mathcal{Y} \end{aligned}$$

4.2. Incorporation of curvature cue

Recognized as the Gestalt law of good continuity, human vision systems have the preference for grouping the smooth contours together. Our method can be extended to encode such curvature prior. The smoothness of contour is traditionally measured by integral of squared curvature of all the contour points. 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

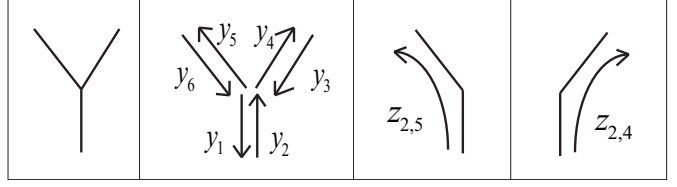


Figure 4. Our junction model. The first figure shows one junction detected in the image. The second figure shows the 6 variables representing the associated edges. The third and fourth figures show two possible L-junctions if edge y_2 is active.

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:

$$E_C(\mathbf{z}) = \alpha_c \sum_{(i,j) \in P_E} u_{ij} z_{ij} \quad (15)$$

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 [11], only the curvature cost within the edge fragments is taken into account. However, our curvature term also penalizes the sharp turns at the junctions.

Our junction model is illustrated in Figure 4. To ensure correct junction configuration, a set of junction constraints are devised, denoted as $\Phi_J(\mathbf{y}, \mathbf{z}) \leq 0$. These constraints are adapted from the connectedness constraints for undirected edges [14]. 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:

$$\sum_{j | (i,j) \in P_E} z_{ij} \geq y_i, \quad \forall i \quad (16)$$

$$z_{ij} \leq y_i, \quad \forall (i, j) \in P_E \quad (17)$$

$$z_{ij} \leq y_j, \quad \forall (i, j) \in P_E \quad (18)$$

The inequities (16) correspond the first part of junction constraint. The inequities (17) (18) correspond to the second part of junction constraints. In sum, our ratio-based segmentation model is as follows:

$$\begin{aligned} \min_{\mathbf{x}, \mathbf{y}, \mathbf{z}} \quad & \frac{E_R(\mathbf{x}) + E_B(\mathbf{y}) + E_C(\mathbf{z})}{A(\mathbf{x})} \quad (19) \\ \text{s.t.} \quad & \Phi_W(\mathbf{x}, \mathbf{y}) = 0 \\ & \Phi_C(\mathbf{y}) = 0 \\ & \Phi_J(\mathbf{y}, \mathbf{z}) \leq 0 \\ & \mathbf{x} \in \mathcal{X}, \mathbf{y} \in \mathcal{Y}, \mathbf{z} \in \mathcal{Z} \end{aligned}$$

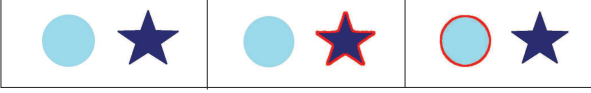


Figure 5. An example in which the curvature term affects our model’s output. The first image is the input image. The second image shows the output without curvature term. Last image shows the output using curvature term. (Better viewed in color.)

The effectiveness of adding curvature term is illustrated by the example in Figure 5. The left is the input image. Two shapes have comparable sizes. The star is favored by the boundary term and the region term due to its stronger contrast to background. Therefore, the model without the curvature term chooses the star as the figure as shown in the middle of Figure 5. When curvature term is added, the smoother disk is extracted as shown in the right of Figure 5.

4.3. Inference by linear relaxation

The energy minimization problems Eq (13) and Eq (19) of our model are nonlinear integer programming. In the following, they are relaxed into linear programs which can be solved in polynomial time. First, the domain of all the labels is relaxed to be interval $[0, 1]$. Second, note that all the terms in the objective function are linear except for E_R which is the sum of the absolute values according to Eq (12). 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 (20). Since the denominator representing the total areas is strictly positive, the fractional program can be transformed into a linear program [6]. In general, the linear fractional program is written as follows:

$$\begin{aligned} \min_{\xi} \quad & \frac{c^T \xi + d}{e^T \xi + f} \\ & A\xi = b \\ & \xi \geq 0 \end{aligned} \quad (20)$$

where the real-valued vector ξ here refers to all the variables, and 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:

$$\begin{aligned} \min_{\eta, \tau} \quad & c^T \eta + d\tau \\ & A\eta = b \\ & \eta \geq 0 \\ & e^T \eta + f\tau = 1 \end{aligned} \quad (21)$$

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

4.4. Implementation details

The boundary gap measure v_i in Eq. (9) equals to the number of edge pixels in the segment minus the sum of the probability of each edge pixel being a true contour point. The probability is estimated according to [11]. The region affinity w_{ij} measure in Eq. (12) is the sum of the similarity 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 pixelwise weight $w(p, q)$ is computed based on the similarity of pixel color and locations using a RBF kernel.

We use the LP_SOLVE library to solve the linear programming problem. An image is usually oversegmented into 300 to 400 superpixels, and our algorithm consists of ten to twenty thousand variables and thirty to forty 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.

5. Extension to other objective functions

Other well established objective function such as normalized cuts can also be transformed into a function based on the edge labels. Although this objective function is more difficult to optimize, it is included in this paper for completeness. The objective function of normalized cuts [18] is defined as:

$$\min_{\mathbf{x} \in \mathcal{X}} \frac{\sum_{ij} w_{ij} |x_i - x_j|}{(\sum_i w_i x_i)(\sum_i w_i (1 - x_i))} \quad (22)$$

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

$$\begin{aligned} \min_{\mathbf{y}} \quad & \frac{\sum_{ij} w_{ij} |m_i^T \mathbf{y} - m_j^T \mathbf{y}|}{(\sum_i w_i m_i^T \mathbf{y})(\sum_i w_i (1 - m_i^T \mathbf{y}))} \\ & s.t. \quad \Phi_C(\mathbf{y}) = 0 \\ & M\mathbf{y} \in \mathcal{X} \\ & \mathbf{y} \in \mathcal{Y} \end{aligned} \quad (23)$$

6. Experiments

To demonstrate the effectiveness of combining region and contour information, our method is compared with the superpixel closure method (SC for short) [11] and the normalized cuts [18] (Ncuts) on the Weizmann horse dataset [5]. Then we show sample results on BSDS 300 dataset. These results show that our method achieves better results than methods only using region or contour cue.

6.1. Comparison with the SC and Ncuts methods

The normalized cuts is a popular segmentation method based on pairwise regional affinities. Similar region cue is

used in our region term. Paper [19] proposes a cost function which is the ratio of contour gap over areas. The SC method, in the standard form, optimize this cost function on superpixels edges. The cost function in [19] is used as the boundary term in our model, and our method used the same set of input superpixels as SC. In this experiment, the curvature term is not used and the weight between the region term and boundary term is fixed by a validation set.

The image set contains salient unoccluded horses in the middle of image. However, obtaining a complete contour of horses is still challenging due to several reasons. For example, there are strong distracting contours in the background and inside the horse region. True contours on the other hand may be faint or missing because of low contrast with the background region. Since the horses usually are not camouflaged, we expect the incorporation of region cue will be helpful in obtaining cleaner contour.

Both SC and our method was initialized on the Pb detection results. The Ncuts implementation is obtained from [18]. The results are shown in Figure 6. SC method outputs ten solutions for each image, the best one is shown in the figure. We can see that the our model’s outputs better separate the horse region from the background. In the SC’s outputs, the legs are often connected as a single blob region. However, the region similarity cue used in our model helps distinguish background region from the foreground horse region. The two-way Ncuts method often cuts out a homogeneous background area. The ten-way segmentation results, however, tend to produce spurious edges (e.g. those in the sky and grass). This results show that region homogeneity cue alone is not enough for segmenting salient foreground region. Note that our results in Figure 6 appear to be a single contour due to the property of the objective function. The winding number constraints, however, do not require the solution to be a Jordan curve.

We qualitatively evaluate our method and SC using the F-measure on 100 horse dataset images. Their model produces up to 10 solutions for each image. According to [11], the F-value of each solution is computed by comparing the segmentation mask with the groundtruth mask. The F-value of an image is the best F-value of all the solutions. The averaged F-value of test set converges to 76.48%. However, as shown in [11], 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 74.12%.

6.2. Tests on BSDS300 dataset

Our method is also tested on the BSDS 300 dataset which contains images of a variety of urban and natural scenes. Our method, which aims to extract salient closed contours is not successful for detecting occluded, obscured, or camouflaged figural objects in this dataset. However, it works very well for images in which the foreground object/region

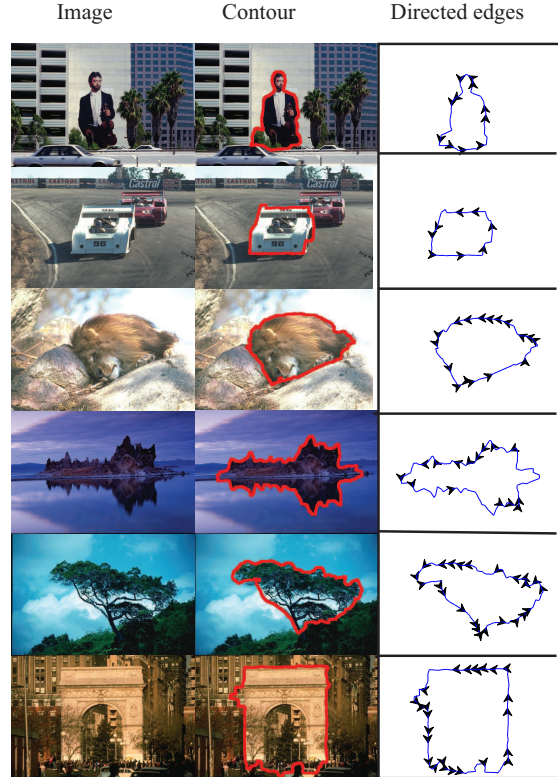


Figure 7. Sample results on BSDS300 dataset. The first column is the input images. The second column is the output contour overlaid on the input images. The third column shows the directed active edges. (Better viewed in color.)

is evident. Some sample results are shown in Figure 7. The first column shows the original image, the second column shows the results of contour overlaid on the image. The third column shows the found directed active edges. These results match our perception of salient region.

7. Conclusion and future work

A winding number based method, is introduced in this paper for enforcing the region-contour consistency constraints. This model is simple and appealing, as it naturally leads to a more compact set of linear constraints, and thus is more efficient than previous methods. Our experiments show that evident improvements can be made for the task of salient contour extraction when both region cue and contour cue are employed. In future, we are interested in finding efficient optimization methods for more complex objective functions and extend this method to multiple-label segmentation.

Acknowledgement. This work is in part funded by Australian Research Council through ARC Discovery grants. NICTA is funded by Australian Government and by ARC.

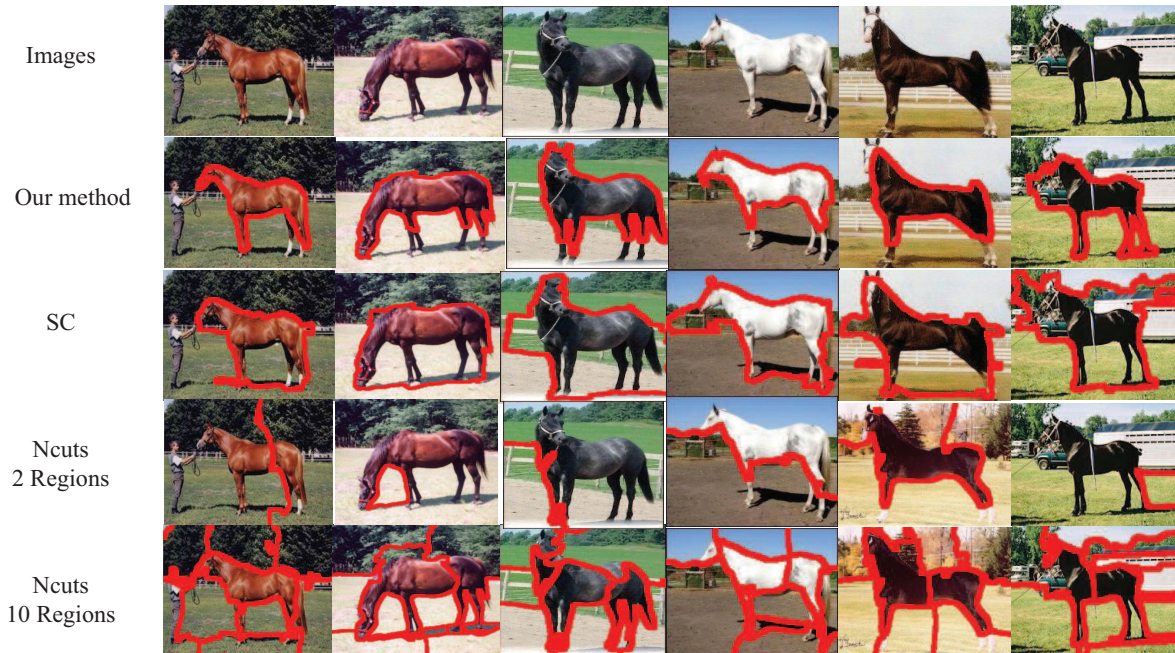


Figure 6. Comparison with superpixel closure method (SC) and the normalized cuts method (Ncuts). The first row shows the input images. The second row shows our results. The third row shows the SC results. Only the best solution of each image is shown out of 10 solutions. The fourth row shows the 2-way segmentation result by Ncuts. The last row shows 10-way segmentation results by Ncuts. (Better viewed in color.)

References

- [1] S. Alpert, M. Galun, A. Brandt, and R. Basri. Image segmentation by probabilistic bottom-up aggregation and cue integration. *IEEE TPAMI*, 34(2):315–327, 2012. 2
- [2] B. Andres, J. H. Kappes, T. Beier, U. Köthe, and F. Hamprecht. Probabilistic image segmentation with closedness constraints. In *Proc. ICCV*, 2011. 1, 2, 3
- [3] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik. Contour detection and hierarchical image segmentation. *IEEE TPAMI*. 2
- [4] C. M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006. 2
- [5] E. Borenstein and S. Ullman. Class-specific, top-down segmentation. In *Proc. ECCV*, pages 109–124, 2002. 6
- [6] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, NY, USA, 2004. 6
- [7] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE TPAMI*, 2001. 1, 2
- [8] J. H. Elder and S. W. Zucker. Computing contour closure. In *Proc. ECCV*, pages 399–412, 1996. 2
- [9] P. F. Felzenszwalb and D. P. Huttenlocher. Efficient graph-based image segmentation. *IJCV*, pages 167–181, 2004. 2
- [10] T. Leung and J. Malik. Contour continuity in region based image segmentation. In *Proc. ECCV*, pages 544–559, 1998. 1, 2
- [11] A. Levinshtein, C. Sminchisescu, and S. Dickinson. Optimal contour closure by superpixel grouping. In *Proc. ECCV*, pages 480–493, 2010. 1, 2, 4, 5, 6, 7
- [12] S. Mahamud, L. R. Williams, K. K. Thornber, and K. L. Xu. Segmentation of multiple salient closed contours from real images. *IEEE TPAMI*, 25(4):433–444, 2003. 2
- [13] D. R. Martin, C. C. Fowlkes, and J. Malik. Learning to detect natural image boundaries using local brightness, color, and texture cues. *IEEE TPAMI*, 26(5):530–549, 2004. 2
- [14] Y. Ming, H. Li, and X. He. Connected contours: A new contour completion model that respects the closure effect. In *Proc. CVPR*, pages 829–836, june 2012. 2, 5
- [15] T. Needham. *Visual Complex Analysis*. Oxford University Press, USA, Feb. 1999. 3
- [16] T. Schoenemann, F. Kahl, S. Masnou, and D. Cremers. A linear framework for region-based image segmentation and inpainting involving curvature penalization. *IJCV*, 99(1):53–68, 2012. 1, 2
- [17] T. Schoenemann, S. Masnou, and D. Cremers. The elastic ratio: Introducing curvature into ratio-based image segmentation. *IEEE TIP*, pages 2565–2581, 2011. 4
- [18] J. B. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE TPAMI*, 22(8):888–905, 2000. 1, 2, 6, 7
- [19] J. S. Stahl and S. Wang. Edge grouping combining boundary and region information. *IEEE TIP*, 16(10):2590–2606, 2007. 1, 2, 4, 7
- [20] B. Sumengen and B. S. Manjunath. Graph partitioning active contours (gpac) for image segmentation. *IEEE TPAMI*, pages 509–521. 1, 2
- [21] M. Tabb and N. Ahuja. Multiscale image segmentation by integrated edge and region detection. *IEEE TIP*, 6(5):642–655, may 1997. 2
- [22] S. Wang, T. Kubota, J. M. Siskind, and J. Wang. Salient closed boundary extraction with ratio contour. *IEEE TPAMI*, 27(4):546–561, 2005. 2, 4
- [23] S. Wang and J. M. Siskind. Image segmentation with ratio cut. *IEEE TPAMI*, 25:675–690, 2003. 1, 2
- [24] L. R. Williams and K. K. Thornber. A comparison of measures for detecting natural shapes in cluttered backgrounds. *IJCV*, 34(2-3):81–96, 1999. 2
- [25] X. Xie and M. Mirmehdi. Rags: Region-aided geometric snake. *IEEE TIP*, pages 640–652, 2004. 1
- [26] S. X. Yu, T. S. Lee, and T. Kanade. A hierarchical markov random field model for figure-ground segregation. *EMMCVPR*, 2001. 2
- [27] L. Zhang and Q. Ji. Image segmentation with a unified graphical model. *IEEE TPAMI*, 32(8):1406–1425, 2010. 1