Kidney segmentation using graph cuts and pixel connectivity

Ashish K. Rudra, Ananda S. Chowdhury, Ahmed Elnakib, Fahmi Khalifa, Ahmed Soliman, Garth Beache, Ayman El-Baz

Department of Electronics and Telecommunication Engineering, Jadavpur University, Kolkata, India
BioImaging Laboratory, Bioengineering Department, University of Louisville, Louisville, KY, USA
Department of Diagnostic Radiology, School of Medicine, University of Louisville, Louisville, KY, USA

A R T I C L E   I N F O
Article history:
Received 23 April 2012
Available online 23 May 2013
Communicated by A. Shokoufandeh

Keywords:
Kidney segmentation
Graph cut
Connectivity analysis
Pixel labeling
Dijkstra’s shortest path algorithm

A B S T R A C T
Kidney segmentation from abdominal MRI data is used as an effective and accurate indicator for renal function in many clinical situations. The goal of this research is to accurately segment kidney from very low contrast MRI data. The present problem becomes challenging mainly due to poor contrast, high noise and partial volume effects introduced during the scanning process. In this paper, we propose a novel kidney segmentation algorithm using graph cuts and pixel connectivity. A connectivity term is introduced in the energy function of the standard graph cut via pixel labeling. Each pixel is assigned a different label based on its probabilities to belong to two different segmentation classes and probabilities of its neighbors to belong to these segmentation classes. The labeling process is formulated according to Dijkstra’s shortest path algorithm. Experimental results yield a (mean ± s.d.) Dice coefficient value of (98.60 ± 0.52)% on 25 datasets.

1. Introduction

Acute rejection – the immunological response of the human immune system to the foreign kidney – is the most important cause of graft failure after renal transplantation (Rigg, 1995). Currently, the diagnosis of rejection is done via biopsy. However, biopsy has the downside effect of subjecting the patients to risks like bleeding and infections (Yang et al., 2001). Therefore, a noninvasive and repeatable technique like computer-aided segmentation of kidney becomes essential for the diagnosis of acute renal rejection. From the accurate segmentation, agent kinetic curves (i.e., signal intensity versus time curves, or perfusion curves) are constructed. From these curves, perfusion-related indexes (e.g., peak signal intensity, time-to-peak, and initial up-slope) are estimated and are used to distinguish between rejection groups and non-rejection groups (Khalifa et al., 2010). Our input for this problem is dynamic MRI data obtained using gradient-echo T1 imaging with a 1.5T MRI scanner (Signa Horizon LX Echo speed; General Electric Medical Systems, Milwaukee, WI, USA). In particular, some images of the time series data have extremely poor contrast. This is the reason why we first focus on kidney segmentation in 2D slices and later show the extension on entire 3D volumes. The main factors which make this problem of kidney segmentation very difficult in addition to extremely poor contrast are unclear borders between the kidney and its background, image acquisition artifacts, image noise, and various pathologies, such as tumors and nephrolithiasis.

Lin et al. (2011) have proposed a solution for fully automatic kidney segmentation. Temporal changes in intensity counts, intensity-pair distribution image contrast enhancement method, adaptive thresholding, and morphological operations are applied for that purpose. However, this work has several processing stages.

In this work, we focus on improving graph cuts, a single-step process, for kidney segmentation. It is a well-known fact that the graph cut-based segmentation method provides globally optimal segmentation results (Boykov and Funka-Lea, 2006; Boykov and Jolly, 2001). Semi-automated kidney segmentation using graph cuts is reported in Hackjoon et al. (2009), Ali et al. (2007). A variety of methods have been proposed to incorporate fixed parametric shape information to spatially constrain the graph min-cut optimization to increase the accuracy of the segmentation. For example, see the works of Slabaugh and Unal (2005), and Freedman and Zhang (2005). Yuksel et al. (2006) use the Poisson distribution and distance maps to compute the shape term of the graph for the segmentation of 2D kidney slices from DCE-MRI. The parametric methods (Ali et al., 2007; Slabaugh and Unal, 2005; Freedman and Zhang, 2005; Yuksel et al., 2006; Malcolm et al., 2007) require a specific shape model which is found to be inadequate for kidney segmentation due to the large inter-patient shape variability. In addition, an interactive initialization is often required to register the shape model with the image. Use of generic shape information...
in the graph cut framework has emerged as another important strategy which obviates the need of computationally intensive registration of shape model and the image. Veksler (2008) used a generic star-shaped prior in graph cut-based segmentation framework. But a disadvantage of Veksler (2008) is that only a shape obeying generic star shape can be extracted. As a statistical alternative, non-parametric shape modeling techniques are introduced (Sabuncu et al., 2009). Freiman et al. (2010) employ an iterative non-parametric model-based graph min-cut approach for kidney segmentation in CT images. There are also several temporal kidney segmentation approaches available in the literature (Chevaillier et al., 2008; Zöllner et al., 2009; Beatrice et al., 2009). Note that the non-parametric as well the temporal methods are computationally quite intensive. So, we notice that both parametric and non-parametric approaches to shape modeling have certain limitations. Moreover, none of the reported methods explicitly address the problem of extremely poor contrast which is the most important concern of this problem. In an earlier work (Rudra et al., 2011), we have shown how the Boykov–Jolly energy function can be modified to handle the problem of low contrast. However, we find that even this method fails to produce very high quality segmentation results for extremely low contrast images. So, we improve the algorithm described in Rudra et al. (2011) by incorporating the connectivity analysis of the image pixels following the Dijkstra’s shortest path algorithm. Note that graph cut provides an efficient global optimization framework where one can integrate region, boundary and connectivity information. Graph cuts using connectivity information can be found in the works of Vincente et al. (2008), where three connectivity constraints are imposed in the segmentation framework. In the proposed work, we use connectivity in an altogether different manner by classifying the pixels into three different groups. The main advantages of our method are as follows: (i) no shape prior is needed, (ii) no additional user interaction is required to perform the connectivity analysis, and (iii) the method can segment very low contrast kidney images with high accuracy in both 2D and 3D.

The rest of the paper is organized in the following manner: in Section 2, we describe the method in details along with necessary theoretical foundations. In Section 3, we present and analyze the experimental results with necessary comparisons. In Section 4 we conclude the paper and mention the directions for future research.

2. Method

An image is modeled as a weighted undirected graph \( G = (V,E) \). Let \( P \) denote the set of all pixels. Each pixel \( p \in P \) constitutes a node/vertex in \( G \). In addition, two special terminal nodes, namely, the ‘source(s)’ and the ‘sink’ (\( t \)) are considered following Ford and Fulkerson, (1962). \( Ne(p) \) denotes the neighborhood of pixel \( p \). The graph \( G \) contains two types of edges: the neighborhood links/n-links (\( N \)) and the terminal links/t-links (\( T \)). Each node \( p \) has two \( t \)-links \( \{s,p\} \) and \( \{p,t\} \), connecting it to the source and the sink. The \( n \)-links are constructed between each pixel \( p \) and its neighboring pixels \( q \ (q \in Ne(p)) \). We use \( |Ne(p)| = 8 \). A new type of edge called connectivity links/c-links (\( C \)) is introduced. The c-links connect every pixel \( p \) to the user specified pixel \( u \), which can be any object seed. The weight of c-link of a pixel depends on how well the pixel is connected to the node \( u \). For this graph \( G \), we can write:

\[
V = P \cup S \cup T \cup U \cup C \tag{1a}
\]

\[
E = T \cup U \cup C \tag{1b}
\]

A. Graph cut for low contrast images

Let \( A \) define segmentation, i.e., classification of all pixels into either “object” or “background”. The corresponding energy function, to be optimized, can be written as:

\[
E(A) = Bo(A) + \lambda R(A) + \mu C(A) \tag{2}
\]

where the boundary properties are given by the term \( Bo(A) \), the region properties are given by the term \( R(A) \), the connectivity properties are given by the term \( C(A) \), and \( \lambda \) and \( \mu \) are weighting factors. For a low contrast image, \( Bo(A) \) and \( R(A) \) can be mathematically expressed as (Rudra et al., 2011):

\[
Bo(A) = \sum_{p,q \in Ne(p)} B_{p,q}(A) \tag{3}
\]

\[
R(A) = \sum_{p \in P} R_p(A_p) \tag{4}
\]

where,

\[
Bo(p,q) = K(p,q) \exp\left(- \left| \frac{1}{2} \right| \right) \cdot \frac{1}{d(p,q)} \tag{5}
\]

\[
R_p(\text{obj}) = H_p^\text{obj} \exp\left(- \left( |S_p - S_b|/D \right) \right) \ln\left| Pr(l_p|\text{obj}) \right| \tag{6}
\]

\[
R_p(\text{bkg}) = H_p^\text{bkg} \exp\left(- \left( |S_p - S_b|/D \right) \right) \ln\left| Pr(l_p|\text{bkg}) \right| \tag{7}
\]

Here, \( Pr(l_p|\text{obj}) \) and \( Pr(l_p|\text{bkg}) \) are obtained from the object and background seeds inputted by the user (Boykov and Funka-Lea, 2006; Boykov and Jolly, 2001). Expressions for \( S_p \) and \( D \) are given below:

\[
S_p = \sum_{i=1}^{N_p} Pr(l_q|\text{obj})/\text{dist}(p,q) \tag{8}
\]

\[
S_b = \sum_{i=1}^{N_p} Pr(l_q|\text{bkg})/\text{dist}(p,q) \tag{9}
\]

\[
D = \sum_{i=1}^{N_p} (1/\text{dist}(p,q)) \tag{10}
\]

The coefficient \( K(p,q) \) in Eq. (5) is evaluated by comparing the probabilities of any two voxels \( p \in P \) and \( q \in Ne(p) \) to belong to the same segmentation class using the terms \( Pr(l_p|\text{obj}) \), \( Pr(l_p|\text{bkg}) \), \( Pr(l_q|\text{obj}) \) and \( Pr(l_q|\text{bkg}) \). The coefficient \( K_p^{\text{obj}} \) in Eq. (6) and the coefficient \( K_p^{\text{bkg}} \) in Eq. (7) are evaluated on the basis of the probability of a single voxel \( p \) to belong to the segmentation class “object” or “background” using the terms \( Pr(l_p|\text{obj}) \) and \( Pr(l_p|\text{bkg}) \). For the detailed algorithm which has been used to evaluate \( K(p,q) \), \( K_p^{\text{obj}} \) and \( K_p^{\text{bkg}} \), see Rudra et al. (2011).

The third term \( C(A) \) in Eq. (2) can be expressed as:

\[
C(A) = \sum_{p \in P} C_p(A_p) \tag{11}
\]

The value of \( C_p \) depends on how well the node \( p \) is connected to the user specified node \( u \). The strength of the connection between \( u \) and \( p \) is determined according to the characteristics of the shortest path (\( W_p \)) between these nodes. The shortest path is obtained following Dijkstra’s shortest path algorithm with an appropriate grammar. Note that this algorithm is applied on a graph \( G_t = G_t(V_t,E_t) \), different from the graph \( G \). For the graph \( G_t \), we can write:

\[
V_1 = P \tag{12a}
\]

\[
E_1 = N \tag{12b}
\]

B. Pixel connectivity analysis

We first introduce some definitions to explain the concept of connectivity.
**Definition 1** (Class A node). A node \( p \) is termed as class A if it has a higher probability to belong to the “object” class, compared to that of the “background” class.

**Definition 2** (Class B node). A node \( p \) is termed as class B if it has higher probability to belong to the “background” class, compared to that of the “object” class.

**Definition 3** (\( L_1 \) node). A node \( p \) is labeled \( L_1 \) if it is a class A node and all intervening nodes on the shortest path \( W_p \) from \( u \) to \( p \) are class A nodes.

**Definition 4** (\( L_2 \) node). A node \( p \) is labeled \( L_2 \) if any of the following conditions hold:

(i) All the intervening nodes on the shortest path \( W_p \) from \( u \) to \( p \) are class A nodes but node \( p \) is a class B node.

(ii) Node \( p \) is a class A node and only one of the intervening nodes of \( W_p \) is a class B node.

**Definition 5** (\( L_3 \) node). A node \( p \) is labeled \( L_3 \) if any of the following conditions hold:

(i) Node \( p \) is a class B node and only one of the intervening nodes of \( W_p \) is a class B node.

(ii) More than one transit on nodes on \( W_p \) are class B nodes.

Let us assume that an user specifies a node \( u \) within the object boundary. So, the node \( u \) is essentially a class A node (vide Definition 1) and it can be easily concluded (vide Definition 3) that its label is \( L_1 \). On the basis of the nature of shortest path between \( u \) and any other node \( p \), we label all the nodes as \( L_1 \) or \( L_2 \) or \( L_3 \). We show the structure of the graph \( G \) we have used in this paper in Fig. 1.

Now, we state and prove two important theorems in the context of labeling of nodes.

**Theorem 1.** For an \( L_1 \) node, at least one of its 8 neighbors must also be an \( L_1 \) node.

**Proof.** Let \( p \) be an \( L_1 \) node. Then, all the intermediate nodes along the shortest path \( W_p \) (from \( u \) to \( p \)) are class A nodes. To determine a shortest path from any node, we only consider a graph with neighborhood edges. So, one of the adjacent nodes of \( p \), say, \( q \), must lie on \( W_p \). Thus, we can write \( q \in \Gamma(W_p) \), where \( \Gamma(W_p) \) is the set of all intermediate nodes along \( W_p \). Since (i) \( q \) is a class A node and (ii) the shortest path \( W_q \) contains only such transit nodes which have higher probability to belong to object class, node \( q \) is an \( L_1 \) node. □

**Theorem 2.** If all the neighbor nodes of a particular node \( p \) are \( L_2 \) or \( L_3 \) nodes, then \( p \) cannot be an \( L_1 \) node.

**Proof.** As all the 8 neighbors \( q \in Ne(p) \) are \( L_2 \) or \( L_3 \) nodes, then all shortest paths from \( u \) to \( q \) must contain at least one class B transit node or the node \( q \) itself must be a class B node. On the other hand, we know that any path from \( u \) to \( p \) must pass through at least one of the \( q \) nodes. So, we can never find any path between \( u \) and \( p \) which is devoid of a class B transit node. Therefore, \( p \) cannot satisfy the second condition of \( L_1 \) node, i.e., along the shortest path from \( u \) to \( p \), all transit nodes should be of type class A. Hence, \( p \) is not an \( L_1 \) node. □
The signed intensity gradient along the path and \( \text{dist}(p, L_s) \) is the distance between the node \( p \) and the closest \( L_1 \) node \( (L_o) \) along the path \( W_p \), in terms of the number of intermediate nodes.

The time-complexity (Cormen et al., 2001) of the connectivity analysis algorithm, applied on the sparse graph \( G(V_1, E_1) \) using binary heap, is \( O(E_1 \log V_1) \). The time-complexity of the Edmonds–Karp enhanced Ford–Fulkerson algorithm, used to find a minimum cut is \( O(VE^2) \). Since, \( E_1 < E \) and \( V_1 < V \), total time-complexity of the proposed method is \( O(E_1 \log V_1) + O(VE^2) = O(VE^2) \).

### 3. Experimental results

We have so far experimented with 25 different T1-weighted MRI datasets with extremely poor contrast. In our protocol, gradient-echo T1 imaging was employed using a 1.5T MRI scanner (Signa Horizon LX Echo speed; General Electric Medical Systems, Milwaukee, WI, USA) with the use of a phased-array torso surface coil. Contrast agent gadoteric acid (Dotarem 0.5 mmol/mL; Guerbet, France) was bolus injected at a rate of 3–4 ml/s, at a dose of 0.2 ml/kgBW. Imaging parameters were: slice thickness = 5 mm, \( TR = 30–40 \) ms, \( TE = 2–3 \) ms, Flip angle = 70°, F.O.V = 38 × 38 cm², and the matrix size 256 × 160. For each patient coronal single slice scan at the level of the renal hilum was taken before contrast. Approximately repeated 80 temporal frames of coronal scans were taken at 3-s intervals after contrast injection.

For experiments in 2D, a single slice (of size 256 × 256 pixels) with extremely poor contrast is chosen from a 3D dataset for all...
the cases. The parameter \( \lambda \) in is experimentally chosen to be 70 and the parameter \( \mu \) is set to 1. The value of the parameter \( \sigma \) in Eq. (5), which can be treated as “camera noise,” is estimated as 3. In our proposed method, the only requirement for the user specified node \( u \) is that it should be within the object. We have compared the performance of our algorithm with two different graph cut-based segmentation methods described in Rudra et al. (2011) and Ali et al. (2007). In Rudra et al. (2011), graph cut was modified to segment brain MRI images suffering from poor contrast. So, we choose to compare the proposed method with that of Rudra et al. (2011) to emphasize the contribution of the connectivity induced energy term. The proposed connectivity-based approach is essentially a shape modeling technique built into the graph cut framework. Therefore, we decide to compare the method of kidney segmentation described in Ali et al. (2007) which employs graph cuts with shape priors. In Figs. 2 and 3, we show two sets of data with inputs (a), segmentation results of Rudra et al. (2011) (b), segmentation results of Ali et al. (2007) (c), and segmentation results of the proposed method (d). The contour of the kidney, obtained from the ground-truth, is shown in red and is superposed on each segmented output. The figures clearly demonstrate that the results in Fig. 1(d) and Fig. 2(d) are much closer in appearance to the respective ground-truths. Dice coefficient is used for quantitative comparison of the performance of these segmentation algorithms. Note that the D.C. is used as a standard measure for evaluating performance of an image segmentation algorithm. Importantly, this is an unbiased measure, i.e., it penalizes both over-segmentation and under-segmentation (Ge et al., 2006). For the dataset shown in Fig. 1(2), respective Dice coefficients using Rudra et al. (2011), Ali et al. (2007) and our method are 68.92% (70.14%), 96.03% (97.18%) and 98.78% (99.13%). Table 3 shows the minimum, mean ± s.d., and maximum Dice coefficient values of the above methods over the 25 datasets. In order to test the statistical significance of our improved accuracy over its competitors, we run the paired \( t \)-tests (Johnson and Bhattacharya, 2006) on the Dice coefficient data. The \( p \)-value obtained from the paired \( t \)-test on Dice coefficient data of the proposed method and that of Ali et al. (2007) is \( 1.47 \times 10^{-16} \). Similarly, the \( p \)-value resulted from the paired \( t \)-test on Dice coefficient data of the proposed method and that of Rudra et al. (2011) is \( 2.02 \times 10^{-33} \). Very low \( p \)-values in both the cases clearly suggest that the proposed method yields statistically significant performance improvement over both (Ali et al., 2007; Rudra et al., 2011). We have also calculated sensitivity and specificity values for our method. The average sensitivity for our method is 99.21% and the average specificity for our method is 99.88%. The proposed method is quite robust in nature both in terms of the sensitivity to the parameters \( \lambda \) and \( \mu \) and sensitivity to the choice of the node ‘\( u \).’ We first show in Fig. 3 that our D.C. changes very little with variations in \( \lambda \) and \( \mu \) in Eq. (2). We next illustrate using Fig. 4 the insensitivity of D.C. to the node ‘\( u \).’ The proposed method on average takes less than 5 s to execute on an

### Table 3

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Min. dice coeff.</th>
<th>Mean (dice coeff.) ± s.d. (dice coeff.)</th>
<th>Max. dice coeff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rudra et al. (2011)</td>
<td>67.13</td>
<td>70.22 ± 1.83</td>
<td>73.22</td>
</tr>
<tr>
<td>Ali et al. (2007)</td>
<td>93.97</td>
<td>96.02 ± 1.13</td>
<td>97.84</td>
</tr>
<tr>
<td>Our method</td>
<td>97.63</td>
<td>98.60 ± 0.52</td>
<td>99.57</td>
</tr>
</tbody>
</table>

Fig. 3. Variations of D.C. with parameters in the energy function. (a) Variations of D.C. vs. \( \lambda \) for various fixed values of \( \mu \). (b) Variations of D.C. vs. \( \mu \) for various fixed values of \( \lambda \).

Fig. 4. Variations of D.C. with different positions of the node \( u \).
Two 3D datasets. In Fig. 5, we show the 3D segmentation results for (98 ± 0.43)% with 99.16% average sensitivity and 99.69% average specificity. In Fig. 5, we plan to apply this method to various other images which exhibit very low contrast like ultrasound images in cardiology, breast cancer detection and study of growth of foetus (Noble and Boukerreo, 2006).

References


