A Shape-Driven MRF Model for the Segmentation of Organs in Medical Images

D.R. Chittajallu S.K. Shah I.A. Kakadiaris

Computational Biomedicine Lab, Depts. of Computer Science, Elec. & Comp. Engineering, and Biomedical Engineering, Univ. of Houston, Houston, TX, USA
{drchittajallu@uh.edu, shah@cs.uh.edu, ioannisk@uh.edu}

Abstract

In this paper, we present a knowledge-driven Markov Random Field (MRF) model for the segmentation of organs in medical images with particular emphasis on the incorporation of shape constraints into the segmentation problem. We cast the problem of image segmentation as the Maximum A Posteriori (MAP) estimation of a Markov Random Field which, in essence, is equivalent to the minimization of the corresponding Gibbs energy function. We then incorporate a set of constraints into the Gibbs energy function that collectively force the resulting segmentation contour/surface to have a shape similar to that of a given shape template. In particular, we introduce a flux-maximization constraint and a generalized template-based star-shape constraint that are encoded into the first- and second-order clique potentials of the Gibbs energy function, respectively. Our main contribution is in the translation of a set of global notions about the shape of the desired segmentation contour into a set of local measures that can be conveniently encoded into the Gibbs energy function and used in combination with other traditionally used constraints derived from image information. In our experiments, we demonstrate the application of the proposed method to the challenging problem of heart segmentation in non-contrast computed tomography (CT) data.

1. Introduction

Image segmentation is, in general, an ill-posed problem and additional constraints must be imposed in order to achieve the desired solution. Commonly used constraints include the traditional regularization constraints and those based solely on image information such as edges. Better solutions to the segmentation problem can be obtained by taking advantage of any prior information available about the class of images being segmented and the objects of interest present in them. In particular, while segmenting organs in medical images, which is our objective in this paper, a significant amount of prior knowledge about the shape, appearance and location of the organs is available that can be used to constrain the solution space of the segmentation problem. However, it is challenging to unify the information from such a wide variety of sources into a single framework.

A variety of approaches have been proposed, both for image segmentation in general [16] and for medical image segmentation in particular [17]. Among the proposed methods, Markov Random Field (MRF) models provide a principled and elegant way to incorporate a wide variety of constraints into the segmentation problem [13]. Despite the attractive properties associated with these models, they remained relatively less popular due to the lack of efficient discrete optimization techniques. Recently, there has been a renewed interest both in discrete optimization in general and in MRF-based models in particular due to the introduction of graph-cuts [2, 9, 11]. Graph-cuts refers to a class of algorithms that use the max-flow/min-cut algorithms to solve discrete optimization problems. There exist certain limitations on the type of energies/models that can be solved using graph-cut based approaches [10]. However, they are far more computationally efficient than existing methods and also provide very attractive guarantees on the optimality of the obtained solution. Recently, there have been a number of efforts to incorporate various types of prior information into MRF-based segmentation models in such a way that they can be minimized using graph-cuts [1, 4, 8, 5, 15, 12, 14]. Among the various types of prior information, the incorporation of shape, in particular, is very challenging. However, if used, it can make the segmentation more robust, particularly when there are neighboring objects with similar appearance, which is often the case in organ segmentation problems. This is precisely the issue that we address here.

Specifically, in this paper, we present a knowledge-driven MRF model for the segmentation of organs in medical images with particular emphasis on the incorporation of shape constraints into the segmentation problem. We cast the problem of image segmentation as the Maximum A Poste-

978-1-4244-6985-7/10/$26.00 ©2010 IEEE

riori (MAP) estimation of a Markov Random Field which, in essence, is equivalent to the minimization of the corresponding Gibbs energy function. We then incorporate a set of constraints into the Gibbs energy function that collectively force the resulting segmentation contour/surface to have a shape similar to that of a given shape template. In particular, we introduce a flux-maximization constraint and a generalized template-based star-shape constraint that are encoded into the first- and second-order clique potentials of the Gibbs energy function, respectively. Central to both of these constraints is the gradient vector field of the shape template's signed distance map, which we refer to as the *shape prior vector field*. The flux-maximization constraint favors a segmentation contour that is locally orthogonal to the shape prior vector field. The template-based star-shape constraint restricts the solution space to a set of contours for which the flux of the shape prior vector field is positive all along the contour. Our main contribution is in the translation of a set of global notions about the shape of the desired segmentation contour into a set of local measures that can be conveniently encoded into the Gibbs energy function and used in combination with other traditionally used constraints derived from image information. Note that even though the proposed shape constraints can be easily extended to the problem of segmenting multiple, even interacting, organs, we limit our discussion in this paper to the single organ (foreground/background) segmentation problem.

This paper is organized as follows: In Section 2, we describe in detail the theory underlying the proposed method. We present in Section 3 an outline of the steps involved in our segmentation algorithm. In Section 4, we demonstrate the application of the proposed method to the challenging problem of heart segmentation in non-contrast CT data and present our segmentation results. Finally, we present our conclusions in Section 5.

2. Theory

In this section, we present the theory underlying the proposed method. Specifically, we begin by formulating the segmentation problem as the minimization of the Gibbs energy function in Section 2.1. We then demonstrate how to incorporate various types of prior information into the Gibbs energy function in Section 2.2. Finally, in Section 2.3, we discuss the minimization of the Gibbs energy function using graph-cuts.

2.1. Formulation of the segmentation problem

The image segmentation problem is a labeling problem. Specifically, consider an image I and let $\mathcal{P} = \{1, 2, ..., M\}$ be the set of M pixels (or voxels) of the image and $\mathcal{L} = \{l_1, ..., l_H\}$ be the set of H labels corresponding to the H objects to be segmented. The goal of image segmentation now is to find a mapping $\phi : \mathcal{P} \rightarrow \mathcal{L}$ that is optimal in some sense. The Markov Random Field (MRF) theory provides an elegant mathematical framework for solving this problem [13]. Using the MRF framework, we define the mapping ϕ as a field $\mathbf{F} = (F_1, ..., F_M)$ of random variables defined on the set of pixels $\mathcal{P}$, where F_i is the random variable associated with the pixel $i \in \mathcal{P}$ and takes on a value from the set of labels $\mathcal{L}$. Any possible assignment of labels to the field of random variables is called a labeling or configuration, which we denote by the vector $\mathbf{f} = (f_1, ..., f_M)$, where f_i is the label assigned to the random variable F_i. Note that every configuration $\mathbf{f}$ defines a segmentation and we denote the set of all possible configurations as $\mathbb{F}$. We also define a neighborhood system $\mathbb{N} = \{\mathcal{N}_i | \forall i \in \mathcal{P}\}$ for the set of pixels $\mathcal{P}$, where $\mathcal{N}_i$ is the set of all neighbors of the pixel $i \in P$. For example, this can be a 4- or 8-neighborhood system for 2D images and a 6- or 26-neighborhood system for 3D images. The random field $\mathbf{F}$ qualifies as an MRF with respect to the neighborhood system $\mathbb{N}$ if and only if it satisfies the following two properties:

$Positivity$:

$$Pr(\mathbf{f}) > 0, \forall \mathbf{f} \in \mathbb{F}$$

$Markovianity$:

$$\Pr(f_i | \{f_j : j \in \mathcal{P} - \{i\}\}) = \Pr(f_i | \{f_j : j \in \mathcal{N}_i\}), \forall i \in \mathcal{P}.$$

Note that, here, we refer to $\Pr(\mathbf{F} = \mathbf{f})$ by $\Pr(\mathbf{f})$ and $\Pr(F_i = f_i)$ by $\Pr(f_i)$. Also, note that the event of assigning a label f_i to a random variable F_i is equivalent to assigning the label f_i to the corresponding pixel $i \in \mathcal{P}$. The Markovian property dictates that the label f_i assigned to a pixel i depends only on the labels assigned to its neighboring pixels defined by the set $\mathcal{N}_i$. This condition is generally true for medical images: the statistics of a pixel in a medical image are related to the statistics of the pixels in a small local neighborhood around it [3]. Now that we have an MRF, we need to find a way to model the probability $\Pr(\mathbf{f}|\mathbf{D})$ of a particular labeling configuration $\mathbf{f}$ given the observed image data $\mathbf{D}$. The Hammersley-Clifford theorem [6] provides an elegant solution to this problem. According to this theorem, an MRF is equivalent to a Gibbs random field (GRF) which includes an interesting and beneficial property: the random field $\mathbf{F}$ qualifies as a Gibbs random field if it obeys the Gibbs distribution that can be specified as follows:

$$\Pr(\mathbf{f}|\mathbf{D}) = Z^{-1} \cdot \exp\left[-E(\mathbf{f}|\mathbf{D})\right], \tag{1}$$

where Z is a normalizing constant and $E(\mathbf{f}|\mathbf{D})$ is the Gibbs energy function. The Gibbs energy function

$$E(\mathbf{f}|\mathbf{D}) = \sum_{c \in \mathbb{C}} V_c(\mathbf{f}_c|\mathbf{D}) \tag{2}$$

is essentially a sum of clique potentials $V_c(\mathbf{f}_c|\mathbf{D})$ over the set $\mathbb{C}$ of all possible cliques. A clique c, in our case, can be defined as a subset of the set $\mathcal{P}$ of pixels such that each member of the set is a neighbor of all the other members. The value of $V_c(\mathbf{f}_c|\mathbf{D})$ depends on the local configuration $\mathbf{f}_c$ of the clique c. The number of pixels in a clique defines the order of the clique and the corresponding clique potential. For the purposes of this work, we only consider first- and second-order cliques. In this case, the Gibbs energy can be expressed as follows:

$$E(\mathbf{f}|\mathbf{D}) = \sum_{i\in\mathcal{P}} V_i(f_i|\mathbf{D}) + \sum_{i\in\mathcal{P}}\sum_{j\in\mathcal{N}_i} V_{ij}(f_i, f_j|\mathbf{D}), \quad (3)$$

where $V_i(f_i|\mathbf{D})$ and $V_{ij}(f_i, f_j|\mathbf{D})$ are the first- and second-order clique potential functions, respectively. The first-order clique potential $V_i(f_i|\mathbf{D})$ in Eq. 3 measures the cost or penalty incurred in assigning a label f_i to the pixel i given the image data $\mathbf{D}$. The second-order clique potential function $V_{ij}(f_i, f_j|\mathbf{D})$ in Eq. 3 measures the cost or penalty incurred in jointly assigning a label f_i to the pixel i and a label f_j to the pixel $j \in \mathcal{N}_i$ given the image data $\mathbf{D}$.

Now that we have found a convenient way to model $\Pr(\mathbf{f}|\mathbf{D})$, the optimal or MAP labeling $\mathbf{f}^*$ of the MRF can be defined as

$$\mathbf{f}^* = \arg\max_{\mathbf{f}\in\mathbb{F}} \Pr(\mathbf{f}|\mathbf{D}), \quad (4)$$

which is equivalent to minimizing the Gibbs energy function $E(\mathbf{f}|\mathbf{D})$ that we will later refer to as the segmentation energy. Note that, in our case, $\mathcal{L} = \{0, 1\}$ because we are concerned with a single organ (foreground/background) segmentation problem. Specifically, the label 0 corresponds to the background object and the label 1 corresponds to the foreground object. In order to achieve the desired segmentation, our next challenges are to define the segmentation energy $E(\mathbf{f}|\mathbf{D})$ and then to find an efficient way to minimize this energy. In Section 2.2 we present our definition of the segmentation energy $E(\mathbf{f}|\mathbf{D})$ and in Section 2.3 we show how to minimize it using graph-cuts.

2.2. Definition of the segmentation energy $E(\mathbf{f}|\mathbf{D})$

We represent the segmentation energy as a combination of three energy functions, each modeling a specific type of prior information, as shown below:

$$E(\mathbf{f}|\mathbf{D}) = E^A(\mathbf{f}|\mathbf{D}) + E^L(\mathbf{f}|\mathbf{D}) + E^S(\mathbf{f}|\mathbf{D}). \quad (5)$$

The energy functions $E^A(\mathbf{f}|\mathbf{D})$, $E^L(\mathbf{f}|\mathbf{D})$ and $E^S(\mathbf{f}|\mathbf{D})$ model prior information about the appearance, location and shape of the organ being segmented, respectively. Each of the three energy functions are in turn expressed as first- or second-order clique potentials or a combination of both.

Our main contribution in this paper is in the definition of the shape prior energy $E^S(\mathbf{f}|\mathbf{D})$. Further details about the definitions of $E^A(\mathbf{f}|\mathbf{D})$, $E^L(\mathbf{f}|\mathbf{D})$ and $E^S(\mathbf{f}|\mathbf{D})$ are available in Sections 2.2.1, 2.2.2, and 2.2.3, respectively.

2.2.1 Appearance prior $E^A(\mathbf{f}|\mathbf{D})$

The energy function $E^A(\mathbf{f}|\mathbf{D})$ in Eq. 5 models prior information about the appearance of the object/organ being segmented. We define $E^A(\mathbf{f}|\mathbf{D})$ as a combination of first- and second-order clique potentials as:

$$E^A(\mathbf{f}|\mathbf{D}) = \sum_{i\in\mathcal{P}} V_i^A(f_i|\mathbf{D}) + \sum_{i\in\mathcal{P}}\sum_{j\in\mathcal{N}_i} V_{ij}^A(f_i, f_j|\mathbf{D}), \quad (6)$$

where $V_i^A(f_i|\mathbf{D})$ and $V_{ij}^A(f_i, f_j|\mathbf{D})$ are the first- and second-order clique potentials modeling regional and boundary appearance information, respectively.

We define the first-order clique potential $V_i^A(f_i|\mathbf{D})$ probabilistically [12] as shown below:

$$V_i^A(f_i|\mathbf{D}) = -log\left[p\left(\mathbf{x} = \mathbf{d}_i; \boldsymbol{\theta}_{f_i}\right)\right], \quad (7)$$

where $p(\mathbf{x}; \boldsymbol{\theta}_{f_i})$ is a probability density function (pdf) of the feature values of the object O_{f_i} associated with the label f_i, $\boldsymbol{\theta}_{f_i}$ is a vector of parameters governing this density function and $\mathbf{d}_i$ is the feature vector corresponding to the pixel i. For most scenarios in organ segmentation, a Gaussian mixture model serves as a good choice for the probability density function. The parameters of the pdf can either be learned *a priori* or estimated from a set of user provided or automatically computed seed regions inside the object.

We define the second-order clique potential $V_{ij}^A(f_i, f_j|\mathbf{D})$ as a piecewise constant prior using a Generalized Potts Interaction model [1] as shown below:

$$\begin{aligned} V_{ij}^A(f_i, f_j|\mathbf{D}) &= K(i, j|\mathbf{D}) \cdot [1 - \delta(|f_i - f_j|)] \\ &= \begin{cases} K(i, j|\mathbf{D}) & \text{if } f_i \neq f_j \\ 0 & \text{otherwise} \end{cases}, \end{aligned} \quad (8)$$

where $K(i, j|\mathbf{D})$ represents the cost associated with object boundaries, that we define as follows:

$$K(i, j|\mathbf{D}) = \exp\left[-\frac{(\mathbf{d}_i - \mathbf{d}_j)^T \boldsymbol{\Sigma}^{-1} (\mathbf{d}_i - \mathbf{d}_j)}{2}\right], \quad (9)$$

where the feature vectors of the pixels i and j are denoted as $\mathbf{d}_i$ and $\mathbf{d}_j$, respectively. The term $\boldsymbol{\Sigma}$ is the covariance matrix that represents the amount of variability allowed between the feature vector values of two neighboring pixels within an object. This function assigns a higher penalty if two neighboring pixels with similar feature vector values

are assigned different labels. Specifically, if the dissimilarity between the two pixels i and j in the feature space is within the amount of variability allowed by $\mathbf{\Sigma}$, then the event of assigning different labels to them is highly penalized.

2.2.2 Location prior $E^L(\mathbf{f}|\mathbf{D})$

The energy function $E^L(\mathbf{f}|\mathbf{D})$ in Eq. 5 models prior information about the location of the objects being segmented. This information is encoded into the first-order clique potential and can be defined as follows:

$$E^L(\mathbf{f}|\mathbf{D}) = \sum_{i \in \mathcal{P}} V_i^L(f_i|\mathbf{D}), \tag{10}$$

where $V_i^L(f_i|\mathbf{D})$ measures the cost of assigning the label f_i to the pixel i given prior information about the location of the object associated with the label f_i. This can be defined in a hard-manner (hard-constraints [1]) based on a set of user defined or automatically computed seed regions that are known to be inside the object. Alternatively, one can define it in a soft manner based on atlas-registration or in terms of fuzzy spatial relationships with neighboring objects [4]. In this paper, we define $V_i^L(f_i|\mathbf{D})$ in a hard-manner as follows:

$$V_i^L(f_i|\mathbf{D}) = \begin{cases} \infty & \text{if } \exists l \in \mathcal{L} - \{f_i\} \text{ s.t. } i \in R_l \\ 0 & \text{otherwise} \end{cases}, \tag{11}$$

where R_l is a seed region that is known *a priori* to be inside the object with label $l \in \mathcal{L}$. In other words, $V_i^L(f_i|\mathbf{D})$ is assigned an infinite cost if pixel i is inside the seed region of any object whose label is not equal to f_i.

2.2.3 Shape prior $E^S(\mathbf{f}|\mathbf{D})$

The energy function $E^S(\mathbf{f}|\mathbf{D})$ in Eq. 5 models prior information about the shape of the object being segmented. Given a shape template of the object/organ being segmented, our objective here is to design a set of constraints that will force the resulting segmentation contour/surface to attain a shape similar to that of the shape template. The biggest challenge here is to translate this global notion of shape similarity into a set of local measures/constraints that can be encoded into the first- or second-order clique potentials of the Gibbs energy function.

We model $E^S(\mathbf{f}|\mathbf{D})$ as a combination of three energy functions, each of them enforcing a specific type of shape constraint, as shown below:

$$E^S(\mathbf{f}|\mathbf{D}) = E^F(\mathbf{f}|\mathbf{D}) + E^C(\mathbf{f}|\mathbf{D}) + E^D(\mathbf{f}|\mathbf{D}), \tag{12}$$

where $E^F(\mathbf{f}|\mathbf{D})$ models the flux-maximization constraint, $E^C(\mathbf{f}|\mathbf{D})$ models the template-based star-shape constraint, and $E^D(\mathbf{f}|\mathbf{D})$ models the distance-to-closest-iso-contour constraint. Central to these three constraints is the gradient vector field of the shape template's signed distance map, which we will refer to as the shape prior vector field and denote by the symbol $\mathbf{t}$. Before computing the shape prior vector field $\mathbf{t}$, the shape template first needs to be aligned to the target image. Further details about the template pose estimation and the definitions of the three shape constraints in Eq. 12 are presented below.

A. Flux-maximization constraint $E^F(\mathbf{f}|\mathbf{D})$

The flux-maximization constraint favors a segmentation contour that is locally orthogonal to the shape prior vector field $\mathbf{t}$. Let S be an arbitrary segmentation contour/surface in the target image space, then the flux of the vector field $\mathbf{t}$ through the segmentation contour S is

$$flux(S) = \int_S \langle \hat{\mathbf{n}}, \mathbf{t} \rangle \, dS, \tag{13}$$

where $\langle , \rangle$ denotes an inner product and $\hat{\mathbf{n}}$ denotes the unit outward normal at each point on the contour S. It can be easily inferred that a contour S has maximal flux if the unit outward normals along the contour align perfectly with the vector field $\mathbf{t}$, which can happen only when the contour S has the same shape as that of the template. Also, among two contours with the same shape, the flux through the contour with a bigger size is greater than the one with a smaller size. This can serve as a good counter force to the shrinking problem of graph-cut based segmentation methods [19]. Note that the concept of flux maximization has been used earlier in image analysis by a few researchers [1, 18]. However, we are the first to thoroughly explore its effectiveness as a template-driven shape-prior within the MAP-MRF pixel labeling formulation of the segmentation problem.

In order to show how to incorporate this constraint into the Gibbs energy function, we will use the divergence theorem for differentiable vector fields that can be stated as:

$$flux(S) = \int_S \langle \hat{\mathbf{n}}, \mathbf{t} \rangle \, dS = \int_R div(\mathbf{t}) \, dR, \tag{14}$$

where R is the region enclosed within the contour S. The divergence theorem states simply that the flux of a vector field through a contour is equal to the integral of its divergence within the region enclosed by the contour. Note that this theorem is applicable under the condition that the associated vector field is differentiable. In our case, it is known that the shape-prior vector field $\mathbf{t}$ is discontinuous at the skeletal points. In order to address this, we smooth the distance map of the aligned shape template before computing

the shape prior vector field $\mathbf{t}$. By virtue of this theorem, it can be easily seen that the flux-maximization constraint can be incorporated into the Gibbs energy function in the form of a first-order clique potential. Based on the above discussion, we define the flux-maximization constraint $E^F(\mathbf{f}|\mathbf{D})$ as shown below:

$$E^F(\mathbf{f}|\mathbf{D}) = \sum_{i \in \mathcal{P}} V_i^F(f_i|\mathbf{D}), \tag{15}$$

where $V_i^F(f_i|\mathbf{D})$ is defined as follows:

$$V_i^F(f_i|\mathbf{D}) = -div(\mathbf{t}_i) \quad \text{if } f_i = 1. \tag{16}$$

B. Template-based star-shape constraint $E^C(\mathbf{f}|\mathbf{D})$

The template-based star-shape constraint restricts the solution space to a set of contours for which the flux of the shape prior vector field $\mathbf{t}$ is positive all along the contour. In other words, it allows only those segmentation contours that are star/convex with respect to the pre-aligned shape template.

In general, a convex shape is one in which, for any point p inside the shape, a line joining p to any other point inside the shape lies entirely inside the shape. A star shape, on the other hand, is defined with respect to a center point [19]. Specifically, an object is said to have a star shape if, for any point p inside the object, the line joining p to the center point lies totally inside the object. We refer to such shapes as being convex/star with respect to a center point. As can be seen, a convex shape is a special case of a star shape, where in any point inside the shape can act as the center point.

We provide an alternative definition to a star shape based on the gradient vector field $\mathbf{q}$ of the signed distance map of its center point. From this point of view, an object is said to have a star shape if the flux of the vector field $\mathbf{q}$ is positive all along the boundary of the object. This definition turns out to be a very interesting way to describe a star shape. Based on this interpretation, we generalize/extend the star shape to a given shape template. Given the shape prior vector field $\mathbf{t}$ of the pre-aligned shape template, an object is said to have a template-based star-shape if the flux of the shape prior vector field $\mathbf{t}$ is positive all along the boundary of the object. Since this is a constraint related to the boundary of the object, it can be incorporated into the Gibbs energy function in the form of a second-order clique potential.

Based on the above discussion, we define the template-based star-shape constraint $E^C(\mathbf{f}|\mathbf{D})$ as shown below:

$$E^C(\mathbf{f}|\mathbf{D}) = \sum_{i \in \mathcal{P}} \sum_{j \in \mathcal{N}_i} V_{ij}^C(f_i, f_j|\mathbf{D}), \tag{17}$$

where $V_{ij}^C(f_i, f_j|\mathbf{D})$ is defined as follows:

$$V_{ij}^C(f_i, f_j|\mathbf{D}) = \begin{cases} \infty & \text{if } (f_i, f_j) = (0,1), \langle \hat{\mathbf{e}}_{ij}, \mathbf{t}_i \rangle > 0 \\ \infty & \text{if } (f_i, f_j) = (1,0), \langle \hat{\mathbf{e}}_{ij}, \mathbf{t}_i \rangle \leq 0 \\ 0 & \text{otherwise} \end{cases},$$

where $\hat{\mathbf{e}}_{ij}$ is a unit vector representing the direction of the edge between pixels i and j. The term $V_{ij}^C(f_i, f_j|\mathbf{D})$ heavily penalizes any part of the object boundary with a negative flux, and hence enforces the template-based star-shape constraint discussed above.

C. Distance-to-closest-iso-contour constraint $E^D(\mathbf{f}|\mathbf{D})$

The energy function $E^D(\mathbf{f}|\mathbf{D})$ softly penalizes any deviation of the segmentation contour from the iso-contour of the shape template that is closest in distance to the true object boundary. However, since the true object boundary is not known, we first solve the segmentation problem using the other two shape constraints. We then compute the iso-contour that is closest to this segmentation result. Next, we penalize any deviation of the segmentation contour from this iso-contour where the cost of deviation is equal to the amount of deviation. Let ρ be the signed distance associated with this iso-contour and let Y denote the signed distance map of the pre-aligned shape template, then we define the energy function in the form of a second-order clique potential as shown below:

$$E^D(\mathbf{f}|\mathbf{D}) = \sum_{i \in \mathcal{P}} \sum_{j \in \mathcal{N}_i} V_{ij}^D(f_i, f_j|\mathbf{D}), \tag{18}$$

where $V_{ij}^D(f_i, f_j|\mathbf{D})$ is defined as follows:

$$V_{ij}^D(f_i, f_j|\mathbf{D}) = \begin{cases} |Y(i) - \rho| & \text{if } (f_i, f_j) = (1,0), \\ & \langle \hat{\mathbf{e}}_{ij}, \mathbf{t}_i \rangle > 0 \\ 0 & \text{otherwise} \end{cases}.$$

D. Template Pose Estimation

Central to all the shape constraints described above is the shape prior vector field $\mathbf{t}$. Before computing $\mathbf{t}$, the shape template must first be aligned to the target image. In order to do this, we first segment the image without any shape constraints and then align the shape template to the resulting segmentation contour/surface using an image registration procedure. This alignment can be achieved in several ways: one common way is to register the binary images of the shape template and the segmentation result obtained without shape constraints. Alternatively, we register the shape prior vector fields of the shape template and the segmentation result obtained without shape constraints under an affine transformation. The registration is performed using the Elastix registration toolkit [7] with an adaptive stochastic gradient descent optimizer and a mutual information metric.

2.3. Minimizing $E(\mathbf{f}|\mathbf{D})$ using graph-cuts

Minimizing the energy function $E(\mathbf{f}|\mathbf{D})$ is a significant part of the challenge in solving an image segmentation problem. Graph-cuts provides an efficient way to optimize such energy functions owing to certain constraints. Particulary in the case of a binary segmentation problem ($\mathcal{L} = \{0, 1\}$), which is our focus in this paper, graph-cuts provides us with a globally optimal solution provided that V_{ij} is a sub-modular function (i.e., $V_{ij}(0,0) + V_{ij}(1,1) \leq V_{ij}(0,1) + V_{ij}(1,0)$) [10]. It can be easily proven that all constraints that we have incorporated are sub-modular in nature, and hence the energy function can be minimized using graph-cuts.

3. Segmentation algorithm

Our segmentation algorithm consists of three stages, and different constraints are put into effect in different stages. In the first stage, we solve the segmentation problem using only appearance (Section 2.2.1) and location priors (Section 2.2.2). In the second stage, we first align the shape template to the segmentation result obtained in the first stage, as described in Section 2.2.3.D. We then solve the segmentation problem by adding in the flux-maximization and template-based star-shape constraints, as described in Sections 2.2.3.A and 2.2.3.B, respectively. Finally, in the third stage, we add the distance-to-closest-iso-contour constraint (Section 2.2.3.C) and solve the segmentation problem using all constraints described. Invariance to template pose can be achieved to an extent by performing multiple iterations of the second and third stages until the segmentation result converges to a steady state. A brief outline of the steps involved in our algorithm is provided below:

Algorithm 1 Outline of proposed segmentation algorithm

Stage-1: Solve the segmentation problem using only appearance and location priors, as described in Sections 2.2.1 and 2.2.2, respectively.

repeat

Stage-2: Align the shape template to the current segmentation result, as described in Section 2.2.3.D. Solve the segmentation problem by adding in the flux-maximization and template-based star-shape constraints, as described in Sections 2.2.3.A and 2.2.3.B, respectively.

Stage-3: Add the distance-to-closest-iso-contour constraint as described in Section 2.2.3.C and solve the segmentation problem using all the constraints.

until convergence is reached

4. Experiments: Heart Segmentation

In this section, we demonstrate how to apply the proposed method to the challenging problem of heart segmentation in non-contrast CT data. Specifically, we use the proposed method to segment the heart in a coronal slice approximately bisecting the ascending aorta (mid-aorta coronal slice). However, in order to use the proposed method, we first need to create a shape template of the heart. We generate it by manually annotating the heart in one of the coronal slices. Figure 1(b) depicts the manually annotated heart in the coronal slice depicted in Figure 1(a). We now describe in detail the steps involved in using the proposed method to segment the heart in a sample test image shown in Figure 1(c).

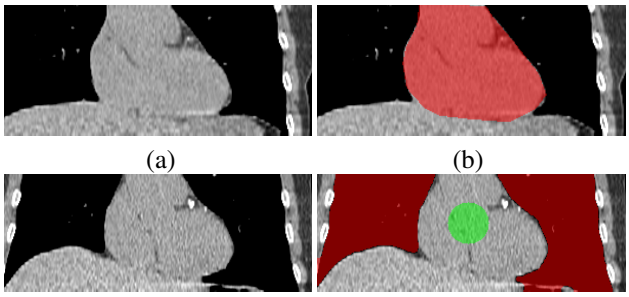

Figure 1. (a) Coronal CT slice used to create the heart template, (b) Manually annotated heart (red) that is used as the heart template, (c) Coronal CT slice of the test image, and (d) Location priors or hard-constraints of the foreground (green) and background (red), respectively.

First, we obtain an initial segmentation using appearance and location priors only. To obtain the location priors, we first segment the lungs using simple thresholding and connected component analysis. We then use the union of the two lung masks to define the location prior of the background. We define the foreground location prior as a small circular region around a point midway between the line joining the centroids of the two lungs. Figure 1(d) depicts the foreground and background location priors computed as described above. We use a Gaussian distribution of the intensity values to define the probability density functions, modeling the regional appearance of the foreground and background objects (Eq. 7). The parameters of these Gaussian distributions are estimated from the respective seed regions depicted in Figure 1(d). Figures 2(a,b) depict the first-order clique potentials V_i^A modeling the regional appearance of the foreground and background objects, respectively. Figures 2(c-f) depict the second-order clique potentials V_{ij}^A (Eq. 8) between neighboring pixels within an 8-neighborhood system. The parameter $\mathbf{\Sigma}$ is defined as the variance of the difference between intensity values of the neighboring pixels within the foreground and background

seed regions depicted in Fig. 1(d). Figure 3(a) depicts the Stage-1 segmentation result obtained using just appearance and location priors as described above.

In the second stage of the algorithm, we first align the heart template to the Stage-1 segmentation result using the procedure described in Section 2.2.3.D. Figure 3(b) depicts the aligned heart template. We then enforce the flux-maximization and the template-based star-shape constraints as described in Sections 2.2.3.A and 2.2.3.B, respectively. Figure 3(c) depicts the segmentation result at the second stage. Note that, the segmentation result confirms to both the shape constraints, but it appears to have a tendency to "jump" across towards larger iso-contours of the shape template in the region inferior to the heart. This problem can be attributed to two reasons: (i) there exists no edge information in this region, and (ii) an iso-contour with a larger perimeter has a higher outward flux. This issue is addressed by the distance-to-closest-iso-contour constraint used in the next stage of our algorithm.

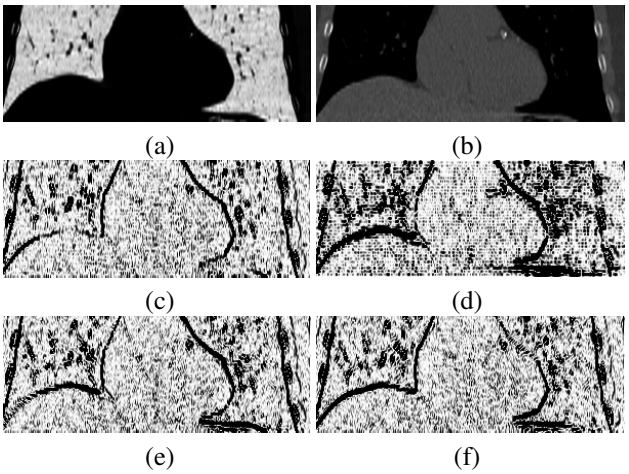

Figure 2. Appearance Priors: (a-b) Depiction of the first-order clique potentials V_i^A (Eq. 7) for foreground and background, respectively, (c-f) Depiction of the second-order clique potentials V_{ij}^A (Eq. 8) between neighboring pixels along the 0° , 90°, 45°, and 135° directions, respectively.

Finally, in the third stage, we add the distance-to-closest-iso-contour constraint as described in Section 2.2.3.C. Figure 3(d) depicts the segmentation result at this stage. As is evident, the distance-to-closest-iso-contour constraint restricts the result to stay close to a single iso-contour and hence addresses the problem faced in stage-2.

We evaluated the proposed method on the mid-aorta coronal slices obtained from non-contrast cardiac CT scans of 12 patients. We randomly selected one of the 12 coronal slices to create the heart template and then used it to segment the heart in all the other slices. We evaluated the accuracy of the segmentation results obtained by measuring the degree

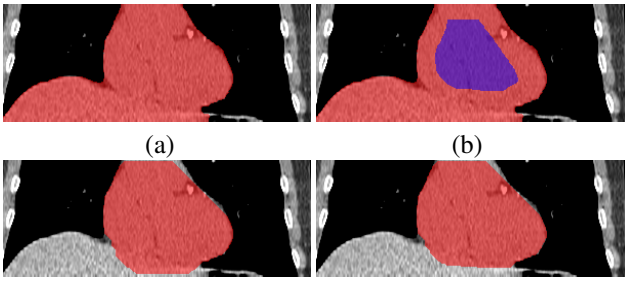

Figure 3. (a) Stage-1: segmentation result obtained using only appearance and location priors, (b) A shrinked version of the shape template (blue) after aligning it to the segmentation result (red) obtained in stage-1, (c) Stage-2: segmentation result obtained by adding in the flux-maximization and the template-based star-shape constraint, (d) Stage-3: segmentation result obtained using all the constraints including the distance-to-closest-iso-contour constraint.

of overlap with manual segmentation performed by an expert where the degree of overlap was estimated using the Dice similarity coefficient (DSC). The mean and standard deviation DSC values were 0.95 and 0.04, respectively. The minimum and maximum DSC values were 0.84 and 0.98, respectively. Figure 4 depicts the segmentation results obtained on six test images from our dataset.

5. Conclusion

In this paper, we have presented a knowledge-driven, MRF model for the segmentation of organs in medical images with particular emphasis on the incorporation of shape constraints into the segmentation problem. Given a shape template of the object/organ, our goal was to design a set of constraints that will force the resulting segmentation contour/surface to attain a shape similar to that of the shape template. To that end, we have explored the incorporation of three types of shape constraints, namely, the flux-maximization constraint, the generalized template-based star-shape constraint and the distance-to-closest-iso-contour constraint. Our main contribution is the translation of a set of global notions about the shape of the desired segmentation contour into a set of local measures that can be conveniently encoded into the Gibbs energy function and used in combination with other traditionally used constraints based on image information. Future work will focus on the incorporation of shape constraints derived from a statistical shape model built from multiple templates.

6. Acknowledgments

This work was supported in part by NSF award CNS-0932272 and the UH Eckhard Pfeiffer Endowment Fund.

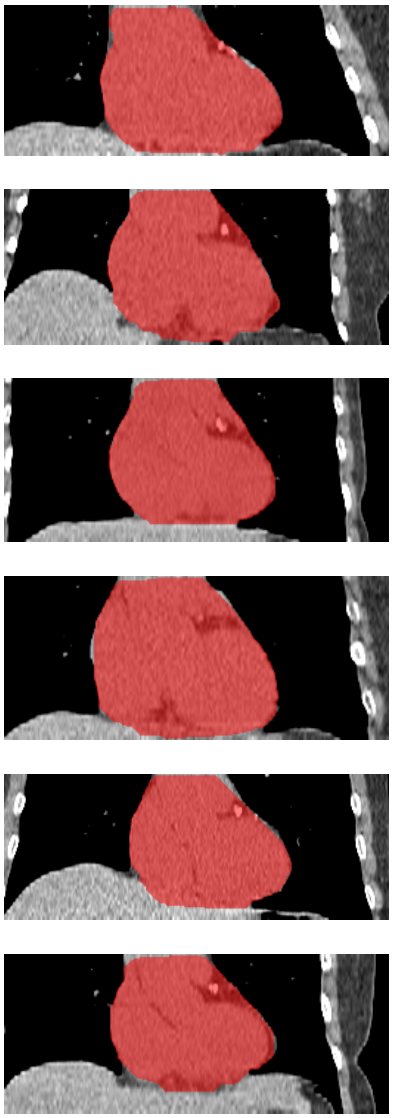

Figure 4. Segmentation results obtained on six test images from our dataset.

Any opinions, findings, conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of our sponsors.

References

[1] Y. Boykov and G. Funka-Lea. Graph cuts and efficient N-D image segmentation. *International Journal of Computer Vision*, 70(2):109–131, Nov. 2006. 1, 3, 4

[2] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1222–1239, Nov. 2001. 1

[3] T. Chen and D. Metaxas. A hybrid framework for 3D medical image segmentation. *Medical Image Analysis*, 9(6):547–565, 2005. 2

[4] D. R. Chittajallu, G. Brunner, U. Kurkure, R. Yalamanchili, and I. A. Kakadiaris. Fuzzy-cuts: A knowledge-driven graph-based method for medical image segmentation. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 715 – 722, Miami Beach, FL, Jun. 20-25 2009. 1, 4

[5] D. Freedman and T. Zhang. Interactive graph cut based segmentation with shape priors. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 755–762, San Diego, CA, Jun. 20-25 2005. 1

[6] J. Hammersley and P. Clifford. Markov field on finite graphs and lattices. Unpublished, 1971. 2

[7] S. Klein, M. Staring, K. Murphy, M. Viergever, and J. Pluim. Elastix: A toolbox for intensity-based medical image registration. *IEEE Transactions on Medical Imaging*, 29(1):196 –205, 2010. 5

[8] V. Kolmogorov and Y. Boykov. What metrics can be approximated by geo-cuts, or global optimization of length/area and flux. In *Proc.* 10^{th} *IEEE International Conference on Computer Vision*, volume 1, pages 564–571, Beijing, China, Oct. 17-20 2005. 1

[9] V. Kolmogorov and C. Rother. Minimizing nonsubmodular functions with graph cuts - a review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(7):1274–1279, 2007. 1

[10] V. Kolmogorov and R. Zabin. What energy functions can be minimized via graph cuts? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(2):147–159, 2004. 1, 6

[11] N. Komodakis and G. Tziritas. Approximate labeling via graph cuts based on linear programming. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(8):1436–1453, 2007. 1

[12] M. Kumar, P. Torr, and A. Zisserman. OBJ CUT. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 18–25, San Diego, CA, Jun. 20-25 2005. 1, 3

[13] S. Li. *Markov random field modeling in computer vision*. Springer-Verlag, 1995. 1, 2

[14] J. Malcolm, Y. Rathi, and A. Tannenbaum. Graph cut segmentation with nonlinear shape priors. In *Proc. IEEE International Conference on Image Processing*, pages 365–368, San Antonio, Texas, Sep. 16-19 2007. 1

[15] V. Nhat and B. Manjunath. Shape prior segmentation of multiple objects with graph cuts. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, Anchorage, AK, 2008. 1

[16] R. Nikhil and K. Sankar. A review on image segmentation techniques. *Pattern Recognition*, 26(9):1277–1294, 1993. 1

[17] D. L. Pham, C. Xu, and J. L. Prince. Current methods in medical image segmentation. *Annual Review of Biomedical Engineering*, 2:315–337, 2000. 1

[18] A. Vasilevskiy and K. Siddiqi. Flux maximizing geometric flows. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(12):1565–1578, 2002. 4

[19] O. Veksler. Star shape prior for graph-cut image segmentation. In *Proc.* 10^{th} *European Conference on Computer Vision*, pages 454 – 467, Marseille, France, Oct. 12-18 2008. 4, 5