Abstract—Image segmentation is, in general, an ill-posed problem and additional constraints need to be imposed in order to achieve the desired segmentation result. While segmenting organs in medical images, which is the topic of 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. Among the various types of prior information, the incorporation of prior information about shape, in particular, is very challenging. In this paper, we present an explicit shape-constrained MAP-MRF-based contour evolution method for the segmentation of organs in 2D medical images. Specifically, we represent the segmentation contour explicitly as a chain of control points. We then cast the segmentation problem as a contour evolution problem, wherein the evolution of the contour is performed by iteratively solving a MAP-MRF labeling problem. The evolution of the contour is governed by three types of prior information, namely: (i) appearance prior, (ii) boundary-edgeness prior, and (iii) shape prior, each of which are incorporated as clique potentials into the MAP-MRF problem. We use the master-slave dual decomposition framework to solve the MAP-MRF labeling problem in each iteration. In our experiments, we demonstrate the application of the proposed method to the challenging problem of heart segmentation in non-contrast computed tomography data.

Index Terms—Medical Image Segmentation, Markov Random Field Model, Contour Evolution, Shape Priors

I. INTRODUCTION

Image segmentation is, in general, an ill-posed problem and additional constraints need to be imposed in order to achieve the desired segmentation result. Commonly used constraints include traditional regularization constraints and constraints based solely on image information (e.g., edges). More accurate solutions to the segmentation problem can be obtained by taking advantage of any prior information available about the class of images being segmented and about the objects of interest present in them. In particular, while segmenting organs in medical images, which is our main concern here, 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. Among the various types of prior information, the incorporation of prior information about shape, in particular, is very challenging. However, if used, it can significantly improve the segmentation results, particularly when there are neighboring objects with similar appearance, which is often the case in organ segmentation problems. In addition, it is even more challenging to unify the information from such a wide variety of sources into a single framework. These are precisely the issues that we intend to address in this paper.

A variety of approaches have been proposed, both for image segmentation in general [1] and for medical image segmentation in particular [2]. Among the proposed methods, MAP-MRF-based models provide a principled and elegant way to incorporate a wide variety of constraints into the segmentation problem [3]. In addition, recent advances in MRF energy minimization algorithms [4], [5], [6], both in terms of computational efficiency and optimality guarantees, have made these models even more attractive. Hence, recently there have been a number of efforts to incorporate different kinds of prior information into MAP-MRF-based segmentation models [7], [8], [9], [10], [11], [12], [13]. With respect to the incorporation of shape priors into the MAP-MRF framework, Freedman et al. [10] used the distance map of the boundary of a pre-aligned shape template to design a pairwise-clique potential that minimizes the deviation between the boundary of the segmented object and the boundary of the shape template. Nhat et al. [11] incorporated shape information in the form of a unary clique-potential that tries to minimize the difference between the Heaviside functions of the segmented object and a given shape template. Chittajallu et al. [12] used the gradient-vector field of the shape template’s signed distance map to design two shape constraints: (i) a flux maximization constraint that maximizes the flux of the vector field through the segmentation contour, and (ii) a template-based star-shape constraint that restricts the class of shapes attainable by the segmentation contour to the ones with a positive flux all along the segmentation contour. Mahapatra et al. [14] used a histogram of the orientations of the lines connecting a pixel to each point on the boundary of a given shape template to incorporate shape information into the unary and pairwise clique potentials of the MRF energy. Each of the MRF-based methods described above formulates the problem of image segmentation as a MAP-MRF pixel labeling problem which we intuitively believe is not naturally suitable for the incorporation of shape priors. The reason behind this intuition, which was also one of the sources of motivation behind pursuing the work described in this paper, is the fact that shape is a property that is associated with the boundary of an object and hence, the most apt formulations for the incorporation...
of shape priors should be the ones that explicitly control the object’s boundary. This is precisely one of the reasons behind the popularity of contour evolution formulations [15], [16], [17], [18] for segmentation based on shape priors. One of the earliest and probably the most significant contribution in this line of approaches was the Active Shape Model (ASM) proposed by Cootes et al. [15] that includes both a shape model and an appearance model, each of which is learned from a training set of images containing the object of interest. In particular, the shape model is built by first discretizing the object’s boundary into a series of landmark or control points followed by the learning of a Gaussian point distribution model of the locations of these landmark points from the training set. The appearance model is built by using the training set to learn a Gaussian model of the intensity gradients along a direction that is normal to the object’s boundary at each landmark point. Leventon et al. [17] and Tsai et al. [18] developed an implicit parametric shape model for level-set based approaches wherein the shape of the object is modeled by learning a Gaussian distribution of the signed distance maps of the object from a set of training images. Prisacariu et al. [19] developed a shape prior for level-set based segmentation and tracking wherein elliptic Fourier descriptors were used to represent the shapes and a non-linear dimensionality reduction technique called Gaussian Process Latent Variable Model was used to learn their lower-dimensional latent space. Even though the contour evolution approaches described above work quite well in practice, they suffer from some drawbacks such as sensitivity to initialization and poor optimality guarantees. In addition, some of these methods require the segmentation energy function to be differentiable, which limits us to an extent in the design of the energy function. Our intention in this paper is to combine the best of both worlds. In other words, we want to take advantage of (i) the inherent flexibility of the MAP-MRF models in accommodating different kinds of prior information accompanied by the recent advances in MRF energy minimization algorithms, and (ii) the ability of contour evolution approaches to explicitly control the boundary of the object which we believe is intuitively more suitable place for the incorporation of shape priors. Note that above we have only reviewed the methods that either bear some similarity to the formulation of the segmentation problem described in this paper or have inspired us in pursuing the work behind this paper. In addition to these, there are a few other ways of localizing objects in images based on shape information, such as the work on deformable template matching as reviewed in [20] and some more recent works [21], [22], [23] in this regard which we have not reviewed here. Also, we would like to point the readers to [24] for a thorough review of the methods available for shape-driven image segmentation.

In this paper, we present an explicit shape-constrained MAP-MRF-based contour evolution method for the segmentation of organs in 2D medical images. Specifically, we represent the segmentation contour explicitly as a chain of control points. We then cast the segmentation problem as a contour evolution problem wherein the evolution of the contour is performed by iteratively solving a MAP-MRF labeling problem [25], [26]. The advantage of this formulation is two-fold. First, the explicit representation of the segmentation contour provides direct control over its evolution. Second, the formulation of the contour evolution problem as a MAP-MRF labeling problem provides a flexible, principled, and elegant way to incorporate various types of prior information into the segmentation problem. In this work, we consider the incorporation of three types of prior information that are commonly available while segmenting organs in medical images:

- **Appearance prior:** Each organ is composed of a set of tissues which in turn have a characteristic appearance or texture in medical images whose knowledge can be used to improve the segmentation. We model this prior information using a generalized probabilistic form of the well-known Chan-Vese functional [27]. Specifically, in this paper, we show how to use the divergence theorem to transform the Chan-Vese functional from a regional integral form into a boundary integral form and then incorporate it as a pairwise-clique potential into the MAP-MRF-based contour evolution formulation of the segmentation problem.

- **Boundary-edgeness prior:** This prior favors the presence of a high intensity gradient along the segmentation boundary. We express this prior in the form of a boundary integral similar to snakes [28] and incorporate it as a pairwise-clique potential into the MAP-MRF-based contour evolution formulation of the segmentation problem.

- **Shape prior:** In general, each organ is characterized by a unique shape and this information, if used, can make the segmentation results more accurate. However, among all the types of prior information available, the incorporation of shape, in particular, is the most challenging. This is because shape is a global property associated with an object in its entirety and all the existing segmentation algorithms are only capable of accommodating local information. Hence, the challenging part in using shape priors is the translation of a set of global notions about the shape of an object into a set of local measures that can be incorporated into the chosen segmentation algorithm. In this paper, we undertake the task of designing a shape constraint for the MAP-MRF-based contour evolution formulation of the segmentation problem. In an earlier work, Besbes et al. [26] proposed a shape prior wherein the distance between the control points of the evolving segmentation contour is constrained to adhere to a distribution learned from the training data. Their shape constraint is invariant to pose, but it is not invariant to scale. Recently, Wang et al. [29] proposed a shape prior that constrains the shapes of the triangles defined by the contour’s control points to adhere to a distribution learned from the training data. The aforementioned shape constraint is both pose- and scale-invariant. However, this shape prior was used within an MRF-based landmark detection formulation of the segmentation problem which did not include the appearance prior and the boundary-
edgeness prior. In this paper, we explore the efficacy of a similar shape constraint in the MAP-MRF-based contour evolution formulation of the segmentation problem.

We use the master-slave dual decomposition framework proposed by Komodakis et al. [30], [31] to solve the MAP-MRF labeling problem in each iteration. Our main contribution is the introduction of a new shape constraint into the MAP-MRF-based explicit contour evolution formulation of the segmentation problem. The shape constraint is theoretically invariant to any similarity transform. Also, by the very nature of its formulation, it largely prevents the evolving contour from self-intersection, which was found to be a common problem with many of the previously proposed explicit contour-evolution formulations of the segmentation problem. In addition to the introduction of a new shape constraint, we also show how to incorporate a regional appearance prior into the MAP-MRF-based explicit contour evolution formulation of the segmentation problem.

This paper is organized as follows: In Section II we describe, in detail, the theory underlying the proposed method. In Section III we present an outline of the steps involved in our segmentation algorithm. In Section IV, we demonstrate the application of the proposed method to the challenging problem of heart segmentation in non-contrast computed tomography (CT) data and then present our segmentation results. Finally, we present our conclusions in Section V.

II. THEORY

In this section, we present the theory underlying the proposed method. Specifically, we begin with a brief review of the formulation of the image segmentation problem as a contour evolution problem in Sec. II-A. Next, we describe how to formulate the contour evolution as a MAP-MRF problem in Sec. II-B. We then demonstrate how to incorporate various types of prior information into the MRF energy function in Secs. II-C - II-G. Finally, in Sec. II-H, we discuss how to minimize the MRF energy using dual decomposition.

A. Image segmentation as a contour evolution problem

In this section, we briefly review the formulation of image segmentation as a contour evolution problem. Let $I$ be a given input image and let $O$ be the object to be segmented. Also, let $C$ be a closed contour in the input image domain $\Omega$, which we will refer to as the segmentation contour. The region $\Omega_{in}$ enclosed within the contour $C$ belongs to the object $O$ and the region $\Omega_{out} = \Omega - \Omega_{in}$ outside the contour $C$ belongs to the background. Let $\mathcal{C}$ be the set of all possible closed contours in the input image domain $\Omega$. The goal of the segmentation problem is to find the contour $C^* \in \mathcal{C}$ representing the boundary of the object to be segmented. The energy minimization formulation is the most common way to model this problem wherein an energy function, $E : \mathcal{C} \rightarrow R$, is defined over the set of all possible contours $\mathcal{C}$ such that its minimum encloses the object of interest as tightly as possible. Following this approach, the segmentation problem gets transformed into the problem of finding the contour $C^* \in \mathcal{C}$ that minimizes the above defined energy function as shown below:

$$C^* = \arg \min_{C \in \mathcal{C}} \{ E(C) \} .$$  (1)

The key factors for achieving the desired segmentation result using the above described energy minimization formulation of the segmentation problem are: (i) the manner in which the segmentation contour $C$ is represented, (ii) the design of the energy function $E(C)$, and (iii) the algorithm used for the optimization of the energy function $E(C)$ over the set of all possible contours $\mathcal{C}$. All these three factors are heavily inter-dependent and have been addressed in various ways in the literature. In this paper, we choose an explicit contour representation wherein the contour is represented explicitly as a chain of control points. We then, cast the contour evolution problem as a MAP-MRF pixel labeling problem and incorporate different kinds of prior information in the form of clique-potentials into the MRF energy. We then use the master-slave dual decomposition framework to minimize the MRF energy.

B. Contour evolution as a MAP-MRF problem

Assuming an explicit representation of the segmentation contour, the contour evolution problem can be cast as a labeling problem. Specifically, let $\mathcal{S} = \{1, 2, ..., M\}$ be the chain of $M$ control points or sites representing the segmentation contour $C$. Let $\mathcal{L} = \{l_1, ..., l_H\}$ be the set of $H$ labels wherein each label $l_i \in \mathcal{L}$ represents a local two-dimensional displacement vector $d_i$ from the quantized deformation space $\mathcal{D} = \{d_1, ..., d_H\}$. The action of assigning a label $l_i \in \mathcal{L}$ to a site $i \in \mathcal{S}$ corresponds to a displacement of $d_i$ of the control point $i$ from its current location in the image domain. The contour evolution problem is essentially a problem of finding a mapping $\phi : \mathcal{S} \rightarrow \mathcal{L}$ that is optimal in some sense. Solving this labeling problem iteratively will result in the evolution of the contour in the image domain. The key challenge now is to find a way to use all the available prior information to constrain the evolution of the contour and guide it towards the desired solution.

The theory of Markov Random Fields provides an elegant mathematical framework for solving such constrained labeling problems [3]. Within this framework, the mapping $\phi$ from the set of sites $\mathcal{S}$ to the set of labels $\mathcal{L}$ is defined using a field $\mathbf{F} = \{F_1, ..., F_M\}$ of random variables wherein each random variable $F_i$ is associated with a site $i \in \mathcal{S}$ and takes on a value from the set of labels $\mathcal{L}$. Any possible assignment of labels to the field $\mathbf{F}$ 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 $F_i$. Now, the labeling problem transforms into the problem of finding a labeling $\mathbf{f}^*$ that maximizes the joint posterior probability $\Pr(\mathbf{f}|\mathbf{D})$ of all the random variables as shown below:

$$\mathbf{f}^* = \arg \max_{\mathbf{f} \in \mathbf{F}} \Pr(\mathbf{f}|\mathbf{D}),$$ (2)
where \( D \) is the observed problem data and \( F \) is the set of all possible labelings of the field of random variables \( F \).

Having defined how to solve the labeling problem, the next challenge is to model the posterior probability \( \Pr(f|D) \) in such a way that we can incorporate all the available prior information into the labeling problem. An elegant solution to this modeling problem emerges if we assume that our random field \( F \) falls into a specific class of random fields called a Markov Random Field (MRF). Specifically, if we assume that our random field \( F \) is an MRF globally conditioned on the observed data \( D \), then by virtue of the Hammersley-Clifford theorem \([3]\) the posterior probability \( \Pr(f|D) \) is given by the Gibbs distribution shown below:

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

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

\[
E(f|D) = \sum_{c \in \chi} V_c(f_c|D)
\]

is essentially a sum of clique potentials \( V_c(f_c|D) \) over the set \( \chi \) of all possible cliques. A clique \( c \), in our case, can be defined as a subset of the set \( S \) of sites such that each member of the set is a neighbor of all the other members wherein two sites \( i \) and \( j \) are considered to be neighbors if the label assigned to the site \( i \) depends on or influences the label assigned to the site \( j \) in some fashion. The notion of neighborhood between two sites is usually specified in the form of a graph \( G = (\mathcal{V}, \mathcal{E}) \) wherein the set of vertices \( \mathcal{V} \) corresponds to the set of sites \( S \) and the set of edges \( \mathcal{E} \) connecting pairs of vertices represents the dependencies between the respective sites. The clique-potential function \( V_c(f_c|D) \) measures the cost of jointly assigning the labels \( f_c \) to the sites in the clique \( c \) given the observed image data \( D \). The number of sites in a clique defines the order of the clique and the corresponding clique potential. For the purposes of this work, we only consider second and third-order clique potentials. In this case, the Gibbs energy can be expressed as follows:

\[
E(f|D) = \sum_{\{i,j\} \in C_2} V_{ij}(f_i, f_j|D) + \sum_{\{i,j,k\} \in C_3} V_{ijk}(f_i, f_j, f_k | D),
\]

where \( C_2 \) and \( C_3 \) are the sets of second- and third-order cliques, respectively. Specifically, in our case, the set \( C_2 \) consists of second-order cliques formed between consecutive sites or control points of the segmentation contour and the set \( C_3 \) contains third-order cliques between selective triplets of control points that are used to encode shape priors as will be described later. Figure 1 contains a schematic depicting the formulation of the contour evolution problem in the form of an MRF labeling problem as described above.

Having modeled the posterior probability \( \Pr(f|D) \) using the Gibbs distribution, the problem of finding a labeling that maximizes the joint posterior probability is equivalent to finding a labeling that minimizes the corresponding Gibbs energy function \( E(f|D) \) which we will henceforth refer to as the segmentation energy or MRF energy. To achieve the desired solution using the above described framework, our next task is to define the segmentation energy \( E(f|D) \) and then find an efficient way to minimize this energy. In Secs. II-C - II-G, we present our definition of the segmentation energy \( E(f|D) \) and describe how to incorporate various types of prior information into it. In Sec. II-H, we show how to minimize it.

C. Definition of the segmentation energy \( E(f|D) \)

We define the segmentation energy \( E(f|D) \) as a combination of four energy functions, each modeling a specific type of prior information, as follows:

\[
E(f|D) = w_A E^A(f|D) + w_B E^B(f|D) + w_R E^R(f|D) + w_S E^S(f|D).
\]

The energy functions \( E^A(f|D) \), \( E^B(f|D) \), \( E^R(f|D) \) and \( E^S(f|D) \) model the appearance prior, boundary edgeness prior, label regularization prior, and shape prior, respectively. The weights \( w_A, w_B, w_R, \) and \( w_S \) are used to tune the relative influence of the respective energy functions. Each of these four energy functions is in turn expressed as a combination of second- and/or third-order clique potentials. Further details about the definitions of \( E^A(f|D) \), \( E^B(f|D) \), \( E^R(f|D) \), and \( E^S(f|D) \) are provided in Secs. II-D, II-E, II-F, and II-G, respectively.

D. Appearance prior \( E^A(f|D) \)

The energy function \( E^A(f|D) \) in Eq. 6 models prior information about the appearance of the object being segmented. An effective way to model this is through a generalized probabilistic form of the Chan-Vese functional \([27]\), which can be expressed as follows:

\[
E_o(C) = \iint_{\Omega_\text{in}} - \log [p_o(I(x,y))] \, dx \, dy + \iint_{\Omega - \Omega_\text{in}} - \log [p_b(I(x,y))] \, dx \, dy + K,
\]

where

\[
\begin{align*}
\Omega_\text{in} & = \{ (x,y) \in \Omega \mid I(x,y) < T \} \\
\Omega - \Omega_\text{in} & = \{ (x,y) \in \Omega \mid I(x,y) \geq T \}
\end{align*}
\]
where \( p_o \) and \( p_b \) are multi-variate probability density functions modeling prior information about the appearance of the object and background regions, respectively, and \( k \) is a constant. As is evident, this is a regional measure and in its current form is not suitable for the explicit contour evolution formulation of the segmentation problem that we have adopted.

To transform Eq. 7 into a boundary-based measure we use the divergence theorem for differentiable vector fields. Specifically, given a two-dimensional differential vector field \( \vec{A}(x, y) = A_x(x, y)\hat{i} + A_y(x, y)\hat{j} \) where \( i \) and \( j \) are unit vectors along the \( x \)- and \( y \)-axis, respectively, the divergence theorem states that the following equality holds:

\[
\int_{\partial \Omega_n} \nabla \cdot \vec{A}(x, y) \, ds = \iint_{\Omega_n} \nabla \cdot \vec{A}(x, y) \, dx \, dy
\]

where \( \nabla \cdot \vec{A}(x, y) = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} \) is the divergence of \( \vec{A} \) and \( \hat{N}(s) \) is the unit outward normal of the contour at \( s \). In other words, the divergence theorem states 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. More importantly, for our work, it shows how to transform a regional measure defined in the region enclosed within a contour into a boundary based measure defined along the contour. To transform the regional measure shown in Eq. 7 into a boundary measure using the divergence theorem shown in Eq. 9, we need to find a vector field \( \vec{A} \) that satisfies the following equation:

\[
\nabla \cdot \vec{A}(x, y) = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} = -\log \left[ \frac{p_o(I(x, y))}{p_b(I(x, y))} \right].
\]

Notice that any vector field \( \vec{A}(x, y) \) which satisfies the equation above will serve our purpose of transforming the regional measure shown in Eq. 7 into a boundary measure.

To facilitate faster computation, we pick the vector field with \( A_x(x, y) = 0 \) and solve Eq. 10 for \( A_x(x, y) \), which yields the following result:

\[
A_x(x, y) = \int_{t=0}^{x} -\log \left[ \frac{p_o(I(t, y))}{p_b(I(t, y))} \right] dt.
\]

Notice that the computation of \( A_x(x, y) \) in the equation above can be performed in constant time by pre-computing an integral image of the measure \( -\log \left[ \frac{p_o(I(x, y))}{p_b(I(x, y))} \right] \) along the \( x \)-direction. Hence, as described above, based on Eqs. 7-10 and by setting \( A_y(x, y) = 0 \), \( E_o(C) \) can be transformed into a boundary based measure shown below:

\[
E_o(C) = \oint C A_x(x(s), y(s)) \cdot \hat{N}(s), ds, \tag{12}
\]

where \( A_x(x, y) \) is computed using Eq. 11. Now that we have a boundary-based measure, we can easily incorporate it into the MRF energy in the form of a second-order clique potential. Based on the above discussion, we define the appearance prior \( E^A(f|D) \) as shown below:

\[
E^A(f|D) = \sum_{(i,j) \in C_2} V_{ij}^A(f_i, f_j|D), \tag{13}
\]

where \( V_{ij}^A(f_i, f_j|D) \) is defined as an integral of the boundary based measure of Eq. 12 along the line segment \( e_{ij} \) connecting the location of the sites \( i \) and \( j \) after the application of the displacements corresponding to the labels \( f_i \) and \( f_j \), respectively, as shown below:

\[
V_{ij}^A(f_i, f_j|D) = \int_{e_{ij}} A_x(x(s), y(s)) \cdot \hat{N}_{e_{ij}}, ds \tag{14}
\]

where \( \hat{N}_{e_{ij}} \) is the unit vector normal to the line segment \( e_{ij} \).

### E. Boundary edgeness prior \( E^B(f|D) \)

The energy function \( E^B(f|D) \) in Eq. 6 models the prior information that an object boundary is more likely to be located in the areas where there is high intensity gradient. Since higher intensity gradient corresponds to a stronger edgeness, we refer to this information as the boundary edgeness prior. We define \( E^B(f|D) \) in the form of a second-order clique potential as shown below:

\[
E^B(f|D) = \sum_{(i,j) \in C_2} V_{ij}^B(f_i, f_j|D), \tag{15}
\]

where \( V_{ij}^B \) is defined as follows:

\[
V_{ij}^B(f_i, f_j|D) = \int_{e_{ij}} \exp \left\{ -\frac{|\nabla I(x(s), y(s))|^2}{2\sigma^2} \right\} ds \tag{16}
\]

where \( |\nabla I(x, y)| \) denotes the gradient magnitude of the image \( I \) at the location \( (x, y) \) and \( \sigma \) is a parameter that can be controlled. Note that that the integral in Eq. 16 is computed along the line segment \( e_{ij} \) connecting the locations of the sites \( i \) and \( j \) after the application of the displacements corresponding to the labels \( f_i \) and \( f_j \). As is evident, the higher the magnitude of the image gradient along the line segment \( e_{ij} \), the lower will be the cost \( V_{ij}^B(f_i, f_j|D) \) of positioning the object boundary at this location.

### F. Label regularization prior \( E^R(f|D) \)

The energy function \( E^R(f|D) \) in Eq. 6 models a simple regularization prior on the labels/displacements assigned to consecutive points on the contour. Specifically, \( E^R(f|D) \) is defined in the form of a second-order clique potential as shown below:

\[
E^R(f|D) = \sum_{(i,j) \in C_2} V_{ij}^R(f_i, f_j|D), \tag{17}
\]

where \( V_{ij}^R \) is defined as follows:

\[
V_{ij}^R(f_i, f_j|D) = ||d_{f_i} - d_{f_j}||_1 \tag{18}
\]

where \( ||\cdot||_1 \) denotes the \( L_1 \)-norm and \( d_{f_i} \in D \) and \( d_{f_j} \in D \) are the displacement vectors corresponding to the labels \( f_i \) and \( f_j \), respectively.
G. Shape prior $E^S(f|D)$

The energy function $E^S(f|D)$ models prior information about the shape of the object being segmented. The shape of an object is the geometric information that remains after filtering out the effects of translation, rotation, and scaling on the object. In other words, two objects embedded in a Euclidean space are said to have the same shape if one can be transformed into the other by a similarity transform (a combination of translation, rotation, and uniform scaling). Based on this notion, we propose a novel way to represent and learn the shape of an object class, which we describe below:

- **Shape representation:** We propose to represent the shape of an object by measuring the angles of the triangles formed by a set of triplets of control points on the object’s boundary. The triplets are selected in such a way that each site is covered in at least one triple-clique. Specifically, given a triplet containing three control points $i$, $j$, and $k$, the shape of the triangle they form can be represented by storing just two of its angles (e.g., $\angle ijk$ and $\angle ikj$) since the sum of three sides of a triangle is equal to 180°. Hence, we represent the shape of the triangle, corresponding to the triplet formed by the control points $i$, $j$, and $k$, by the two element descriptor $s_{ijk} = [\angle ijk, \angle ikj]$. Note that the above described representation is invariant to any similarity transform.

- **Shape learning:** Given a specific object class, even though the global shape remains the same, the presence of a certain amount of intra-object variability is natural. Using the above described shape representation, we can easily capture the intra-object shape variability by learning the distributions of the angles of the triangles of all the triplets from a set of training images. Specifically, given a triplet formed by the control points $i$, $j$, and $k$, we will learn the distribution $p_{ijk}(s_{ijk})$ where $s_{ijk}$ is the shape descriptor described above. We model $p_{ijk}$ using a gaussian distribution.

Using the above shape model, we now define the energy function $E^S(f|D)$ in the form of a third-order clique potential as shown below:

$$E^S(f|D) = \sum_{i,j,k} V^S_{ijk}(f_i, f_j, f_k|D),$$  \hspace{1cm} (19)

where $V^S_{ijk}(f_i, f_j, f_k|D)$ is defined as follows:

$$V^S_{ijk}(f_i, f_j, f_k|D) = -\log(p_{ijk}(s_{ijk}))$$  \hspace{1cm} (20)

where $s_{ijk}$ is the shape descriptor of the triangle formed by the triplet of the three control points $i$, $j$, and $k$ after the application of the displacements corresponding to the labels $f_i$, $f_j$, and $f_k$, respectively, and $p_{ijk}$ is the learned probability density function of the shape descriptor corresponding to this triplet.

Ideally, we would want to enforce the above described shape-constraint on the set of all possible triplets of control points. However, this would make the minimization of the MRF energy computationally very expensive since the number of third-order cliques in the MRF would be $O(n^3)$, where $n$ is the number of control points used to represent the segmentation contour. Hence, we need to minimize the number of selected triplets as much as possible without compromising much on the effectiveness of the shape constraint. Also, note that each site or control point must be a part of at least one triplet. Otherwise, some of the sites will not be directly influenced by the shape constraint and might act against it as the contour evolves in the image space.

H. Minimization of the segmentation energy $E(f|D)$

Having formulated the contour evolution problem as an MRF energy minimization problem, our next task is to find a way to minimize the MRF energy function $E(f|D)$. In our case, the MRF energy is a multi-label MRF energy with higher-order clique potentials, which makes its minimization an NP-hard problem. We use the general master-slave dual decomposition framework proposed by Komodakis et al. [30], [31] to minimize our MRF energy. The key design decision to make while using this framework is finding an efficient way to decompose the original MRF, called the master MRF, into a set of slave MRFs that are easier to optimize. To this end, we first convert the neighborhood graph associated with our MRF into a factor graph representation. We then decompose the factor graph into as many factor trees as there are triple-cliques in such a way that each triple-factor belongs to one and only one factor tree. To generate such a factor tree, we first start with an empty tree. We then pick a triple-factor and add it to the tree. Next, we add all the pairwise factors defined between consecutive control points of the contour. This results in a graph with three cycles. We now remove one pairwise factor from each cycle resulting in a factor tree. We then repeat the same procedure for each triple-factor. Figure 2 illustrates this process for an exemplary elliptical contour with two third-order cliques. Specifically, Figures 2(a) and (b) depict the clique graph and the corresponding factor graph of the elliptical contour. Figures 2(c) and (d) depict the factor trees of the slave MRFs corresponding to two third-order cliques which were generated as described above. Inference in each of the factor trees can be performed exactly and very efficiently using the max-product belief propagation algorithm [32], [33].
III. SEGMENTATION ALGORITHM

A brief outline of the steps involved in our segmentation algorithm is given below:

Algorithm
1. Initialize the segmentation contour
2. Compute the appearance prior $E^A(f \mid D)$ (Sec. II-D)
3. Compute the boundary edgeness prior $E^B(f \mid D)$ (Sec. II-E)
4. Compute the label regularization prior $E^L(f \mid D)$ (Sec. II-F)
5. Compute the shape prior $E^S(f \mid D)$ (Sec. II-G)
6. Minimize the segmentation energy $E(f \mid D)$ (Sec. II-H)
7. Halt, if the Hausdorff distance between the previous and the current segmentation contour is less than a pre-defined threshold. Else, use the current contour for initialization in the next iteration and go to Step 2.

IV. RESULTS

In this section, we present the results obtained using the proposed method on both synthetic and real data. Specifically, in Section IV-A we demonstrate the efficacy of the proposed shape constraint (Sec. II-G) in recovering the learned shape upon initialization with a randomly perturbed initial contour. In Section IV-B, we demonstrate the application of the proposed method to the challenging problem of heart segmentation in non-contrast computed tomography (CT) data and then present our segmentation results.

A. Shape recovery experiment

To investigate the efficacy of the proposed shape constraint (Sec. II-G), we synthetically generated an elliptical contour represented by 12 control points as depicted in Figure 3(a). Next, we considered the fully connected set of all possible third-order cliques between the 12 control points as depicted in Figure 3(b). We then learned the shape of triangles corresponding to each of the third-order cliques as described in Sec. II-G. Starting from a randomly perturbed initial contour depicted in Figure 3(c), and performing contour evolution using shape prior alone, we were able to perfectly recover the shape of the elliptical contour (Figure 3(d)).

Fig. 3. Shape Recovery Experiment: (a) elliptical contour used to learn the shape model, (b) fully connected set of third-order cliques, (c) randomly perturbed initial contour, and (d) final shape recovered by performing contour evolution using only the shape prior.

B. Heart segmentation

Heart segmentation is an essential step in a variety of applications concerning the diagnosis and treatment of cardiovascular disease (CVD). However, the segmentation of heart in non-contrast CT data, in particular, is very difficult due to the presence of neighboring organs that are very similar in appearance to heart and due to the lack of sufficient edge information at the interface between the heart and these organs. Most of the methods proposed in the literature for the segmentation of the heart were either focused on contrast-enhanced CT data (wherein the contrast with neighboring organs is better) or on other imaging modalities [34]. However, these methods are not directly applicable to non-contrast CT data, and major adaptations are needed. In addition, majority of the previous work on heart segmentation is focused on the segmentation of its internal structures (particularly the left ventricle) whereas, in our research, we are particularly interested in the segmentation of the heart as a whole. To the best of our knowledge, Moreno et al. [35], [36] were the first to propose a method for the segmentation of the heart as a whole in non-contrast CT data. They introduced a fuzzy theoretic model that captures anatomical knowledge of the heart’s location and used it both to define a region of interest, and to drive the evolution of a deformable model to segment the heart within this region. Later, Isgum et al. [37] proposed a method based on atlas-based registration for the segmentation of the heart and the aorta in non-contrast CT data. They performed a non-rigid registration of multiple atlases to the target image and obtained the segmentation result by combining the decisions from all the registered atlas images using spatially varying decision fusion weights that are derived from local assessments of registration quality. Recently, we introduced a shape-driven MRF model for the 2D heart segmentation problem [12] wherein the gradient vector field of the signed distance map of a given shape template was used to incorporate a set of shape constraints into the MAP-MRF pixel labeling formulation of the segmentation problem.

In this section, we demonstrate the application of the proposed method to the problem of 2D 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, which we will henceforth refer to as the mid-aorta coronal slice. Our dataset consists of mid-aorta coronal slices taken from 76 non-contrast EBCT scans, out of which 36 were randomly selected for learning the proposed statistical shape model and the remaining 40 were used for testing the performance of our algorithm.

To apply the proposed method to the problem of heart segmentation, we first need to choose a graph structure for the MRF. We represent the boundary of the heart by a chain of 20 equally spaced control points which constitute the sites of our MRF. This results in 20 second-order cliques defined between consecutive points of the heart contour. Next, we need to choose the set of third-order cliques for the shape constraint. Choosing the set of all possible triplets of 20 points would make the minimization of the MRF energy computationally intractable. Hence, we need to reduce the number of triplets as much as possible but we should also keep in mind that each of the 20 sites must be a member of at least one triplet for the shape-constraint to be effective. Based on these considerations, we generate a small subset of the set of all possible triplets wherein we form the triplets by connecting every pair of...
consecutive points on the heart contour to the point that is farthest in distance from their midpoint. This way each and every site is a member of at least two third-order cliques. Also, every pair of consecutive contour points is a member of one third-order clique. Figure 4(b) depicts the MRF clique-graph constructed in the above described fashion.

The next step is to learn the shape model as described in Sec. II-G. Given the manually annotated heart contours in all the training images, we pick one of them as a reference shape and align all the other heart contours to this reference shape by registering their signed distance maps using a similarity transform. We then find the corresponding points in all the aligned heart contours for each of the 20 control points. Once the point correspondences have been established, we learn the shape of the triangles corresponding to each of the selected third-order cliques as described in Sec. II-G. Figure 4(a) depicts the reference image overlayed with a manually annotated heart mask.

Given a test image, our first step is to find a good initialization for the segmentation contour. We achieve this by registering the image corresponding to the reference shape with the given test image using a similarity transform. Next, we define the label space corresponding to the local displacement vectors by periodically sampling eight local displacement vectors from two concentric rings around the location of the control point. We then proceed as described in Steps 2-6 of our segmentation algorithm (Sec. III). Figure 4(c) depicts a sample test image overlayed with the initial segmentation contour which was computed as described above. Figure 4(d) depicts the automatically computed foreground and background seed regions in the test image. These regions are used to compute the foreground and background appearance models that are used to compute the appearance prior (see Eq. 7). Figures 4(e) and (f) depict the probability images of the foreground and background, respectively. Figure 4(g) depicts the final segmentation result obtained using the proposed method and figure 4(i) depicts the manually annotated heart mask which is considered as the groundtruth. Figure 4(h) depicts the segmentation result obtained using the proposed method without the shape prior wherein the segmentation contour got trapped in a local minimum. For the aforementioned experiments, the weights $w_A, w_B, w_R,$ and $w_S$ in Eq. 6 were set to 1.5, 1.0, 0.2, and 40, respectively. These weights were determined experimentally and discovering a way to automatically estimate them from training data is a part of our future work. The proposed method takes around 4 minutes to segment an image from our dataset on an Intel i7 2.6 GHz computer with 12 GB of RAM. A significant decrease in computational time can be achieved by adopting a multi-resolution approach for contour evolution. Most of our algorithm was implemented in MATLAB.

We evaluated the segmentation accuracy of the proposed method on mid-aorta coronal slices of 40 patients by comparing the segmentation results obtained using the proposed method against manual annotations performed by an expert. In particular, the accuracy of the segmentation results was quantified using two well known metrics, namely: (i) Dice similarity coefficient (DSC), and (ii) Hausdorff distance. We also compared the results obtained using the proposed method with our previous work [12] on this problem, the standard Active Shape Model (ASM) proposed by Cootes et al. [15], and the results obtained by using the proposed method without the shape prior. In Table I, we present statistics of the DSC and Hausdorff distance obtained using the proposed method.
the standard ASM, our previous work on this problem, and the proposed method without the shape prior. As is evident from the results, the proposed method yields significantly better results in the presence of the shape prior. It also outperforms the standard ASM method in terms of both, the DSC and the Hausdorff distance. The results obtained using our previous work [12] are similar to the proposed method in terms of the DSC. However, with respect to the Hausdorff distance the proposed method produces slightly better results.

Figure 5 depicts a comparison of the results obtained using the proposed method with the two aforementioned methods and the groundtruth on a sample test image from our dataset. Figure 6 depicts the segmentation results obtained using the proposed method on six test images from our dataset.

V. Conclusion

In this paper, we have presented an explicit shape-constrained MAP-MRF-based contour evolution method for 2D image segmentation. Specifically, we represent the contour explicitly as a chain of control points. We then cast the segmentation problem as a contour evolution problem wherein the evolution of the contour is performed by iteratively solving a MAP-MRF labeling problem. The contour evolution is governed by three types of prior information, namely: (i) appearance prior, (ii) boundary-edgeness prior, and (iii) shape prior, each of which is incorporated as clique potentials into the MAP-MRF model. Our main contribution is in the introduction of a new shape constraint into the MAP-MRF explicit contour evolution formulation of the segmentation problem. As the contour evolves we constrain the shape of the triangles defined by the contour’s control points to adhere to a multivariate normal distribution learned from the training data. In practice, the proposed shape constraint largely prevents the evolving contour from self-intersection, which was found to be a common problem with many of the previously proposed explicit contour-evolution formulations of the segmentation problem. Also, theoretically, the proposed shape constraint is invariant to any similarity transform. However, the suboptimal nature of the MRF energy minimization algorithm requires us to initialize the contour closer to the target pose to achieve the desired segmentation result. In addition to the introduction of a new shape constraint, we also show how to incorporate a regional appearance prior into the MAP-MRF-based explicit contour evolution formulation of the segmentation problem. Our future work will focus on extending the proposed method to 3D and on analyzing the training data to automatically compute a minimal subset of the set of all possible triplets that is effective in constraining the shape of the evolving contour while keeping the MRF energy minimization computationally feasible.

VI. Acknowledgements

This research was funded in part by the National Science Foundation (CPS-0932272) and the Hugh Roy and Lillie Cranz Cullen foundation.

REFERENCES

Fig. 6. Heart Segmentation results obtained using the proposed method on selected images from our dataset.


