Graph Cuts with Invariant Object-Interaction Priors: Application to Intervertebral Disc Segmentation

Ismail Ben Ayed[1,2], Kumaradevan Punithakumar[1,2], Gregory Garvin[2], Walter Romano[2], and Shuo Li[1,2]

[1] GE Healthcare, London, ON, Canada
[2] The University of Western Ontario, ON, Canada

Abstract. This study investigates novel *object-interaction* priors for graph cut image segmentation with application to intervertebral disc delineation in magnetic resonance (MR) lumbar spine images. The algorithm optimizes an original cost function which constrains the solution with learned prior knowledge about the geometric interactions between different objects in the image. Based on a global measure of similarity between distributions, the proposed priors are intrinsically invariant with respect to translation and rotation. We further introduce a scale variable from which we derive an original *fixed-point equation (FPE)*, thereby achieving scale-invariance with only few fast computations. The proposed priors relax the need of costly pose estimation (or registration) procedures and large training sets (we used a single subject for training), and can tolerate shape deformations, unlike template-based priors. Our formulation leads to an *NP-hard* problem which does not afford a form directly amenable to graph cut optimization. We proceeded to a relaxation of the problem via an *auxiliary function*, thereby obtaining a nearly real-time solution with few graph cuts. Quantitative evaluations over 60 intervertebral discs acquired from 10 subjects demonstrated that the proposed algorithm yields a high correlation with independent manual segmentations by an expert. We further demonstrate experimentally the invariance of the proposed geometric attributes. This supports the fact that a single subject is sufficient for training our algorithm, and confirms the relevance of the proposed priors to disc segmentation.

1 Introduction

Accurate segmentation of lumbar spine discs in magnetic resonance (MR) images is useful in quantifying intervertebral disc degeneration (IDD) and assisting surgical spine procedures [10]. Quantitative disc measurements often resort to time-consuming, manual segmentations [12]. Related works generally focused on automating vertebrae segmentation [7,8,4], and only few studies investigated disc segmentation [10,5,15] or detection [6,14]. The problem is acknowledged difficult because of the similarities in intensity and shape profiles between the discs and their surrounding regions (refer to the example in Fig. 2 b). Based on standard

G. Székely and H.K. Hahn (Eds.): IPMI 2011, LNCS 6801, pp. 221–232, 2011.

techniques such as fuzzy clustering [10], watersheds [5], and edge detection [15], existing methods require intensive user inputs (e.g. user landmarks for each disc [10]), a large training set, pose registration, and a heavy computational load. The detection algorithms in [6,14] used probabilistic graphical models to embed prior information on the geometric interaction between pairs of discs. Enforcing the distance between the discs to fall within a known range, these algorithms led to promising detection results. Unfortunately, they yield only disc centroids, not segmentations. Furthermore, their interaction priors are not scale-invariant.

In the scope of image segmentation, embedding priors on the geometric interactions between different objects has been generally avoided, mainly because it leads to challenging optimization problems. Existing geometric priors commonly bias each single target region towards a known set of *template* shapes or atlases, independently of other related regions [10,13,17]. Although very useful in some cases, template-based priors require costly pose estimation (or registration), as well as a large set of training examples. Furthermore, they are sensitive to shape deformations. An unknown pathological case outside the set of learned templates, for instance a degenerated disc with shape irregularities, may not be recovered.

A recent notable study by Toshev et al. [16] demonstrated that a *global* shape description can yield very competitive image segmentation results. The shape model of an object is the histogram of distances and orientations corresponding to all pairs of points on the training shape, and the segmentation is sought following the optimization of the L_1 similarity between histograms, an *NP-hard* problem which the authors solve via semidefinite programming relaxation [16]. Unlike template-based priors, this global shape description is translation invariant, requires a single training example, and can tolerate shape deformations. Unfortunately, the description is not scale or rotation invariant and, therefore, requires heavy computations to handle scale variations. It is worth noting that the contribution in [16] follows on several recent segmentation studies which have shown the usefulness of intensity (or color) priors based on global measures of similarity between distributions [2,11,1]. Although helpful, such intensity priors are not sufficient to obtain satisfying segmentations when the target regions and their surrounding structures have almost the same intensity profiles (refer to the example in Fig. 2 b).

Inspired by the global shape and intensity descriptions in [16,2,11,1], we propose novel *object-interaction* priors and their application to intervertebral disc delineation in MR lumbar spine images. The algorithm optimizes an original cost function which constrains the solution with learned prior knowledge about the geometric interactions between different objects in the image. Based on the Bhattacharyya measure and the distributions of the geometric relationships between pairs of points within different objects, the proposed priors are intrinsically invariant with respect to translation and rotation. We further introduce a scale variable from which we derive an original *fixed-point equation (FPE)*, thereby achieving scale-invariance with only few fast computations. The proposed priors relax the need of costly pose estimation (or registration) procedures and large

training sets (we used a single subject for training), and can tolerate shape deformations, unlike template-based priors. Our formulation leads to an *NP-hard* problem which does not afford a form directly amenable to efficient graph cut optimization. We proceeded to a relaxation of the problem via an *auxiliary function*, thereby obtaining a nearly real-time solution with few graph cuts. Quantitative evaluations over 60 intervertebral discs acquired from 10 subjects demonstrated that the proposed algorithm yields a high correlation with independent manual segmentations by an expert. We further demonstrate experimentally that the proposed geometric attributes do not vary significantly from one subject to another (refer to the illustration in Fig. 4). This experiment supports the fact that a single subject is sufficient for training, and confirms the relevance of the proposed priors to disc segmentation.

2 Formulation

Consider a MR spine image function $I(p) = I_p : \mathcal{P} \subset \mathbb{R}^2 \rightarrow \mathcal{I} \subset \mathbb{R}$, with $\mathcal{P}$ the image domain and $\mathcal{I}$ the set of intensity variables. Given a simple user input, which consists of an elliptic approximation of the boundary of only one disc, the purpose is to delineate all subsequent discs in the image (cf. the examples in Fig. 3). For each disc, the solution is efficiently obtained following the minimization of an original *discrete* energy containing three distribution similarity measures. The first measure is an intensity prior which embeds information about image data within the target disc. The last two measures are geometric priors which embed information about the interactions between neighboring discs.

2.1 The Energy

General definitions and notations: To introduce the energy, we first consider the following general definitions for any labeling (or segmentation) $\mathcal{L}(p) = \mathcal{L}_p : \mathcal{P} \rightarrow \{0, 1\}$, any function $J(p) = J_p : \mathcal{P} \rightarrow \mathcal{J}$, and any set of variables $\mathcal{J} \subset \mathbb{R}$.

- $\mathbf{R}_1^{\mathcal{L}}$ and $\mathbf{R}_0^{\mathcal{L}}$ are the complementary regions defined by $\mathbf{R}_1^{\mathcal{L}} = \{p \in \mathcal{P} / \mathcal{L}_p = 1\}$ and $\mathbf{R}_0^{\mathcal{L}} = \{p \in \mathcal{P} / \mathcal{L}_p = 0\} = \mathcal{P} \setminus \mathbf{R}_1^{\mathcal{L}}$.
- $\mathbf{P}_{\mathcal{L},J}$ is the kernel density estimate (KDE) of the distribution of function $J : \mathcal{P} \rightarrow \mathcal{J}$ within region $\mathbf{R}_1^{\mathcal{L}}$:

$$\forall j \in \mathcal{J}, \quad \mathbf{P}_{\mathcal{L},J}(j) = \frac{\sum_{p \in \mathbf{R}_1^{\mathcal{L}}} K(j - J_p)}{\mathbf{A}(\mathbf{R}_1^{\mathcal{L}})} \text{ with } K(y) = \frac{1}{\sqrt{2\pi\sigma^2}} exp^{-\frac{(y)^2}{2\sigma^2}} \quad (1)$$

$\mathbf{A}(\mathbf{R})$ denotes the number of pixels within region $\mathbf{R}$, and K is the Gaussian kernel (σ is the width of the kernel).

- $\mathcal{B}_{\mathcal{J}}(f, g)$ is the *Bhattacharyya* coefficient measuring the amount of overlap (similarity) between two distributions f and g: $\mathcal{B}_{\mathcal{J}}(f, g) = \sum_{j \in \mathcal{J}} \sqrt{f(j)g(j)}$. Note that the values of $\mathcal{B}_{\mathcal{J}}$ are always in $[0, 1]$, where 0 indicates that there is no overlap, and 1 indicates a perfect match between the distributions.

The intensity prior: Let $\mathcal{M}_I$ denotes a model distribution of intensity learned from image data I within the user-provided elliptic region (cf. the examples in Fig. 3). The minimization of this prior identifies each of the subsequent discs as a region whose intensity distribution most closely matches $\mathcal{M}_I$:

$$\mathbf{B}_I(\mathcal{L}) = -\mathcal{B}_{\mathcal{I}}(\mathbf{P}_{\mathcal{L},I}, \mathcal{M}_I) = -\sum_{i \in \mathcal{I}} \sqrt{\mathbf{P}_{\mathcal{L},I}(i)\mathcal{M}_I(i)} \tag{2}$$

Although helpful, this prior is not sufficient to obtain satisfying segmentations because the discs and some surrounding structures have almost the same intensity profiles (Fig. 2 b depicts an example of segmentation with this prior).

The object-interaction priors: Assume an elliptic approximation of the boundary of a previously segmented disc is given. Let $O \in \mathcal{P}$ be the center of the ellipse and $\boldsymbol{u}$ a unit vector pointing along the minor axis (an illustration is given in Fig. 1). For each point $p \in \mathcal{P}$, let $\boldsymbol{dp}$ the vector pointing from O towards p. Consider the following geometric functions:

$$\begin{cases} C(p) = C_p = \frac{\langle \boldsymbol{dp}, \boldsymbol{u} \rangle}{\|\boldsymbol{dp}\|} = cos(\alpha_p) : \mathcal{P} \rightarrow \mathcal{C} \subset [-1, 1] \\ D(p) = D_p = \|p - O\| : \mathcal{P} \rightarrow \mathcal{D} \subset \mathbb{R} \end{cases} \tag{3}$$

C measures the cosine of the angle between vectors $\boldsymbol{dp}$ and $\boldsymbol{u}$, whereas D evaluates the distance between p and O. To constrain the segmentation with prior geometric information, we propose to optimize the following two constraints.

1. The angle-distribution prior: We assume that the distribution of angle function C within the target disc follows a model $\mathcal{M}_{\mathbf{C}}$ which can be learned from a different training subject. To find a disc region whose angle-function distribution most closely matches $\mathcal{M}_{\mathbf{C}}$, we propose to minimize:

$$\mathbf{B}_C(\mathcal{L}) = -\mathcal{B}_{\mathcal{C}}(\mathbf{P}_{\mathcal{L},C}, \mathcal{M}_{\mathbf{C}}) = -\sum_{c \in \mathcal{C}} \sqrt{\mathbf{P}_{\mathcal{L},C}(c)\mathcal{M}_{\mathbf{C}}(c)} \tag{4}$$

This geometric prior is invariant to translation, rotation, and scale of a pair of discs. We examined such invariance experimentally: using manual segmentations, we plotted in the first line of Fig. 4 the angle-function distributions corresponding to 10 different subjects and 3 different disc pairs. The high similarity between these 10 distributions supports the fact that a single subject is sufficient for training, and confirms the invariance of the prior.

2. The distance-distribution prior: We assume that the distribution of distance function D within the target disc follows a model $\mathcal{M}_D$ learned from the training subject. Our purpose is to find a region whose distance distribution most closely matches $\mathcal{M}_D$ by minimizing

$$\mathbf{B}_D(\mathcal{L}) = -\mathcal{B}_{\mathcal{D}}(\mathbf{P}_{\mathcal{L},D}, \mathcal{M}_D) = -\sum_{d \in \mathcal{D}} \sqrt{\mathbf{P}_{\mathcal{L},D}(d)\mathcal{M}_D(d)} \tag{5}$$

This geometric prior is invariant to translation and rotation, but not to scale. Using manual segmentations, we plotted in the second line of Fig. 4 the distance distributions corresponding to 10 different subjects and 3 different disc pairs. The distributions have similar shapes, but different supports. The shifts between the distributions are due to inter-subject variations in scale.

Introducing a scale variable: To achieve scale invariance of the distance prior, we relax the assumption that the distribution of D within the target disc follows exactly the learned model $\mathcal{M}_D$. We rather assume that it belongs to the following set of distributions parametrized with a scale variable s:

$$\{\mathcal{M}_D(.,s) : \mathcal{D} \times \mathbb{R} \rightarrow [0,1] / \mathcal{M}_D(d,s) = \mathcal{M}_D(d+s), s \in \mathbb{R}\} \tag{6}$$

In this case, we rewrite the distance-distribution prior as follows

$$\mathbf{B}_D(\mathcal{L}, s) = -\mathcal{B}_{\mathcal{D}}(\mathbf{P}_{\mathcal{L},D}, \mathcal{M}_D(.,s)) = -\sum_{d \in \mathcal{D}} \sqrt{\mathbf{P}_{\mathcal{L},D}(d)\mathcal{M}_D(d+s)} \tag{7}$$

where s is an additional variable which has to be optimized with the labeling. Based on global rather pixel-wise information, the proposed geometric priors

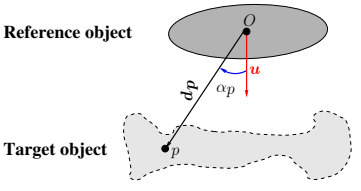


Fig. 1. The geometric relationships between pairs of points within different objects: the upper object (reference) is segmented with the proposed method at a previous stage. The lower object is the target sought at the current stage of the segmentation process.

relaxes (1) extensive learning/modeling of geometric characteristics (we use a single subject for training) and (2) complex optimization with respect to several translation and rotation parameters.

We propose to minimize an energy containing the intensity and object-interaction priors as well as a regularization term for smooth boundaries:

$$\{\mathcal{L}^{opt}, s^{opt}\} = \arg\min_{\mathcal{L},s} \mathcal{F}(\mathcal{L}, s) = \underbrace{\mathbf{B}_I(\mathcal{L})}_{Image\ \ prior} + \underbrace{\beta\left(\mathbf{B}_C(\mathcal{L}) + \mathbf{B}_D(\mathcal{L}, s)\right)}_{Object-interaction\ \ priors} + \underbrace{\lambda \mathbf{S}(\mathcal{L})}_{Smoothness} \tag{8}$$

$\mathbf{S}(\mathcal{L})$ ensures label consistency of neighboring pixels ($\mathcal{N}$ is a neighborhood system): $\mathbf{S}(\mathcal{L}) = \sum_{\{p,q\} \in \mathcal{N}} \frac{\delta_{\mathcal{L}(p) \neq \mathcal{L}(q)}}{\|p-q\|}$ with $\delta_{x \neq y} = 1$ if $x \neq y$ and $\delta_{x \neq y} = 0$ if $x = y$. γ and λ are positive constants balancing the contribution of each term.

2.2 Optimization

Energy (8) depends on two type of variables (labeling $\mathcal{L}$ and scale variable s). Therefore, we proceed to an iterative two-step optimization strategy. The first step consists of fixing s and optimizing $\mathcal{F}(\mathcal{L}, s)$ with respect to the labeling via auxiliary-function relaxation and graph cuts. The second step consists of finding the optimal scale variable via fixed-point-equation updates, given the labeling provided by the first step. The algorithm iterates these two steps until convergence. Each step decreases $\mathcal{F}(\mathcal{L}, s)$ with respect to a variable. Thus, the algorithm is guaranteed to converge.

Step 1–Graph Cut optimization via auxiliary-function relaxation: The *global* terms $\mathbf{B}_I$, $\mathbf{B}_C$ and $\mathbf{B}_D$ in (8) are not directly amenable to max-flow optimization because they do not reference pixel or pixel-neighborhood penalties. They evaluates a global similarity measure between distributions and, therefore, the ensuing optimization problem is challenging and *NP-hard*. To obtain a solution efficiently, we proceed to an *auxiliary-function relaxation* of the problem in (8). Rather than optimizing directly cost function $\mathcal{F}$, the relaxation optimizes iteratively a sequence of upper bounds of $\mathcal{F}(\mathcal{L}, s)$, denoted $\mathcal{A}(\mathcal{L}, \mathcal{L}^n, s)$, $n = 1 \ldots$:

$$\mathcal{L}^{n+1} \;=\; \arg \min_{\mathcal{L} \in \{0,1\}} \; \mathcal{A}(\mathcal{L}, \mathcal{L}^n, s)\,, \quad n \;=\; 1 \ldots \tag{9}$$

under the following constraints:

$$\begin{cases} \mathcal{F}(\mathcal{L}, s) \;\leq\; \mathcal{A}(\mathcal{L}, \mathcal{L}^n, s)\,, \quad n \;=\; 1 \ldots \\ \mathcal{F}(\mathcal{L}, s) = \mathcal{A}(\mathcal{L}, \mathcal{L}, s) \end{cases} \tag{10}$$

$\mathcal{A}$ is called auxiliary function of $\mathcal{F}$. Such relaxations are commonly used in the Nonnegative Matrix Factorization (NMF) literature for challenging problems [9]. Using the constraints in (10), and by definition of minimum in (9), one can show that the sequence of solutions in (9) yields a monotonically decreasing sequence of $\mathcal{F}$:

$$\mathcal{F}(\mathcal{L}^{(n)}, s) = \mathcal{A}(\mathcal{L}^{(n)}, \mathcal{L}^{(n)}, s) \geq \mathcal{A}(\mathcal{L}^{(n+1)}, \mathcal{L}^{(n)}, s) \geq \mathcal{F}(\mathcal{L}^{(n+1)}, s) \tag{11}$$

Furthermore, $\mathcal{F}(\mathcal{L}^{(n)}, s)$ is lower bounded because the Bhattacharyya measures are upper bounded by one. Therefore, $\mathcal{F}(\mathcal{L}^{(n)}, s)$ converges to a minimum of $\mathcal{F}$, and the solution is obtained by the optimal labeling at convergence.

Auxiliary function of the proposed energy: To introduce an auxiliary function of the proposed energy, let us first consider the following proposition:

Proposition 1: *Given a fixed labeling $\mathcal{L}^n$, for any labeling $\mathcal{L}$ verifying $\mathbf{R}_1^{\mathcal{L}} \subset \mathbf{R}_1^{\mathcal{L}^n}$ and $\forall \alpha \in [0,1]$, we have the following upper bound of the proposed energy:*

$$\begin{aligned} \mathcal{F}(\mathcal{L}, s) \;\leq\; & \mathcal{A}(\mathcal{L}, \mathcal{L}^n, s, \alpha) = \mathcal{A}_I(\mathcal{L}, \mathcal{L}^n, 0, \alpha) + \beta(\mathcal{A}_C(\mathcal{L}, \mathcal{L}^n, 0, \alpha) + \mathcal{A}_D(\mathcal{L}, \mathcal{L}^n, s, \alpha)) \\ & + \lambda \mathbf{S}(\mathcal{L}) \end{aligned} \tag{12}$$

where $\mathcal{A}_J(\mathcal{L}, \mathcal{L}^n, s, \alpha)$ has the following general form for any function $J \in \{I : \mathcal{P} \to \mathcal{I}, C : \mathcal{P} \to \mathcal{C}, D : \mathcal{P} \to \mathcal{D}\}$:

$$\mathcal{A}_J(\mathcal{L}, \mathcal{L}^n, s, \alpha) = \sum_{p \in \mathbf{R}_0^{\mathcal{L}}} m_{p,J,s}(\mathbf{0}) + (1-\alpha) \sum_{p \in \mathbf{R}_1^{\mathcal{L}}} m_{p,J,s}(\mathbf{1}), \tag{13}$$

with $m_{p,J,s}(\mathbf{0})$ and $m_{p,J,s}(\mathbf{1})$ given for each p in $\mathcal{P}$ by:

$$\begin{cases} m_{p,J,s}(\mathbf{0}) = \frac{\mathcal{L}_p^n}{\mathbf{A}(\mathbf{R}_1^{\mathcal{L}^n})}\left(\mathbf{B}_J(\mathcal{L}^n,s) + \sum_{j\in\mathcal{J}} K(j-J_p)\sqrt{\frac{\mathcal{M}_J(j+s)}{\mathbf{P}_{\mathcal{L}^n,J}(j)}}\right) \\ m_{p,J,s}(\mathbf{1}) = \frac{\mathbf{B}_J(\mathcal{L}^n,s)}{\mathbf{A}(\mathbf{R}_1^{\mathcal{L}^n})} \;\text{ with }\; \mathbf{B}_J(\mathcal{L},s) = -\sum_{j\in\mathcal{J}}\sqrt{\mathbf{P}_{\mathcal{L},J}(j)\mathcal{M}_J(j+s)} \end{cases} \tag{14}$$

To prove proposition 1, we apply the same principle steps of the proof we detailed recently in [1] to each of the Bhattacharyya constraints in cost function (8). We omit the details here due to space limit.

One can further verify that, for $\alpha = 0$, $\mathcal{A}(\mathcal{L},\mathcal{L},s,\alpha) = \mathcal{F}(\mathcal{L},s)$, i.e., the bound in (12) is an auxiliary function of the proposed energy. This instructs us to consider the following procedure to optimize $\mathcal{F}$ over the labeling:

- $n = 0$; *Initialize the labeling* $\mathcal{L}^0 = \mathcal{L}_0$*; Initialize* α*:* $\alpha = \alpha_0$ *with* $0 < \alpha_0 < 1$
- *Repeat the following two steps until convergence*
  1. $\mathcal{L}^{(n+1)} = \arg\min_{\mathcal{L}:\mathbf{R}_1^{\mathcal{L}} \subset \mathbf{R}_1^{\mathcal{L}^n}} \mathcal{A}(\mathcal{L},\mathcal{L}^n,s,\alpha)$
  2. *Decrease* α*:* $\alpha = \alpha^\rho$ *with* $\rho > 1$

Graph Cuts: Now notice that the auxiliary function $\mathcal{A}(\mathcal{L},\mathcal{L}^n,s,\alpha)$ in *step 1* of the above procedure has the form of the sum of *unary* and *pairwise (sub-modular)* penalties, unlike the initial energy $\mathcal{F}$. In combinatorial optimization, a global optimum of such form can be computed efficiently in low-order polynomial time via a graph cut [3]. The graph-cut (or max-flow) algorithm of Boykov and Kolmogorov is well established in the computer vision literature. Therefore, we omit the details of this algorithm here, and refer the reader to [3].

Step 2–Fixed-point-equation (FPE) updates of the scale variable: With the labeling fixed, this step optimizes $\mathcal{F}$ with respect to s. Following the variable change $d \leftarrow d - s$, the derivative of $\mathcal{F}$ with respect to s reads:

$$\frac{\partial\mathcal{F}}{\partial s} = -\beta\frac{\partial\sum_{d\in\mathcal{D}}\sqrt{\mathbf{P}_{\mathcal{L},D}(d)\mathcal{M}_D(d+s)}}{\partial s} = -\beta\sum_{d\in\mathcal{D}}\frac{\partial\mathbf{P}_{\mathcal{L},D}(d-s)}{\partial s}\sqrt{\frac{\mathcal{M}_D(d)}{2\mathbf{P}_{\mathcal{L},D}(d-s)}} \tag{15}$$

Using the KDE expression in (1), we also have:

$$\frac{\partial\mathbf{P}_{\mathcal{L},D}(d-s)}{\partial s} = \frac{\sum_{p\in\mathbf{R}_1^{\mathcal{L}}}\frac{\partial K(d-s-D_p)}{\partial s}}{\mathbf{A}(\mathbf{R}_1^{\mathcal{L}})} = \frac{\sum_{p\in\mathbf{R}_1^{\mathcal{L}}}(d-s-D_p)K(d-s-D_p)}{\sigma^2\mathbf{A}(\mathbf{R}_1^{\mathcal{L}})} \tag{16}$$

Embedding this derivative in (15), setting the obtained expression equal to zero, and after some algebraic manipulations, the necessary condition for a minimum of $\mathcal{F}$ with respect to s can be expressed as the following fixed-point equation:

$$s - g(s) = 0 \text{ where } g(s) = \frac{\sum_{d\in\mathcal{D}}\sum_{p\in\mathbf{R}_1^{\mathcal{L}}}(d-D_p)K(d-D_p-s)\sqrt{\frac{\mathcal{M}_D(d)}{\mathbf{P}_{\mathcal{L},D}(d-s)}}}{\sum_{d\in\mathcal{D}}\sum_{p\in\mathbf{R}_1^{\mathcal{L}}}K(d-D_p-s)\sqrt{\frac{\mathcal{M}_D(d)}{\mathbf{P}_{\mathcal{L},D}(d-s)}}} \tag{17}$$

Therefore, the solution of (17) can be obtained by the following *fixed-point-equation updates*:

$$s^{n+1} = g(s^n),\ n = 1, 2, \ldots \tag{18}$$

Let s^{opt} be the limit of sequence s^n at convergence. We have:

$$s^{opt} = lim_{n\to+\infty} s^{n+1} = lim_{n\to+\infty} g(s^n) = g(lim_{n\to+\infty} s^n) = g(s^{opt}) \tag{19}$$

Consequently, s_{opt} is a solution of the necessary condition obtained in (17).

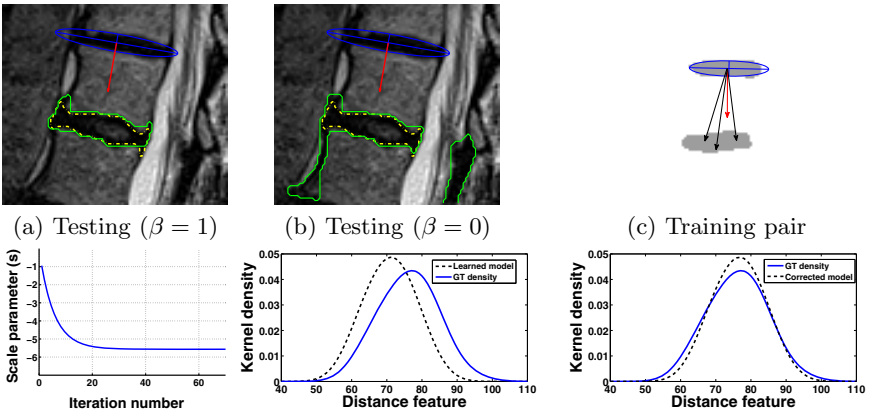
2.3 Experiments

The evaluation was carried out over 10 midsagittal T2-weighted MR lumbar spine images[1]. We proceeded to a leave-one-out approach: for each testing case, we used a single different subject for training. A total of 60 lumbar discs were automatically delineated, and the results were compared to independent manual segmentations by an expert. Although we focus on lumbar spine discs here (6 discs per subject) as in [6], the formulation can be readily extended to the whole spine without additional user effort.

In the following, we first describe a typical example which illustrates explicitly the effect of the proposed object-interaction priors, the effect of the fixed-point equation we derived, and the robustness of the algorithm with respect to pose variations. Then, we describe a quantitative and comparative performance analysis using several accuracy measures. Finally, we give a representative sample of the results, and examine experimentally the invariance of the proposed geometric attributes, i.e., the angle and distance distributions, over ground-truth segmentations of 10 subjects and several disc pairs.

A Typical example: The continuous green curve in Fig. 2 (b) depicts the segmentation boundary obtained *without* the object-interaction priors. This solution, which corresponds to $\beta = 0$ in problem (8), included erroneously some neighboring structures in the final disc region (the ground truth is depicted with the discontinuous yellow curve). On the contrary, *with* the proposed priors, the algorithm yielded a solution very close to the ground truth (Fig. 2 a). The blue curve depicts an elliptical approximation of a previously segmented disc. Fig. 2 (c) depicts the disc pair in the sole training subject, and illustrates how the algorithm handles successfully the differences in pose (translation, rotation, and scale) between the training and testing pairs. The second line in Fig. 2 illustrates the effect of the fixed-point computations corresponding to the same example. (d) shows the fast convergence of such computations: the optimal scale variable s^{opt} is typically obtained within less than 30 iterations. The discontinuous line in Fig. 2 (e) depicts the distance model $\mathcal{M}_D$ learned from the training pair in (c), whereas the continuous one depicts the distance distribution corresponding to the ground truth in the test image in (a). The shift between these two distributions is due to the difference in scale between the training and testing subjects.

[1] 2D T2-weighted MR spine images are commonly used in clinical practice thanks to their short acquisition time and ability to depict disc degeneration [10].



(a) Testing ($\beta = 1$) (b) Testing ($\beta = 0$) (c) Training pair

(d) $s^{opt} = -5.56$ (e) learned model $\mathcal{M}_D$ (f) corrected model $\mathcal{M}_D(., s^{opt})$

Fig. 2. A typical example. The green curves in (a) and (b) depict respectively the segmentation obtained with and without the object-interaction priors. The yellow discontinuous curve in (a)-(b) depicts the ground truth. (c) shows the disc pair of the training subject. (d) shows the fast convergence of the proposed fixed-point-equation computations: s^{opt} is obtained within less than 30 iterations (1.38 sec). (e) and (f) show respectively the learned model ($\mathcal{M}_D$) and the model corrected with s^{opt} ($\mathcal{M}_D(., s^{opt})$), both displayed with the ground-truth distance distribution (GT density). The number of graph cut iterations is 9 (3.3 sec). $\lambda = 2.5 \times 10^{-4}$. The distributions were estimated using 192 bins and a kernel width $\sigma = 15$. *The proposed formulation handles intrinsically translation and rotation variations, compute efficiently the optimal scale, and does not enforce a systematic bias towards the shape of the training discs.*

Fig. 2 (f) shows that $\mathcal{M}_D(., s^{opt})$, i.e., the model corrected with the optimal scale variable s^{opt}, befits much better the ground-truth distribution. The overall computation time is 4.68 sec (The graph cuts took 3.3 sec and the fixed-point iterations 1.38 sec). It is worth noting that the algorithm did not bias the solution towards the shape of the training discs in (c), unlike template-based priors.

Visual inspection: Fig. 3 depicts a representative sample of the results for the lumbar spines of three subjects. The green curves depict the segmentations obtained with the proposed algorithm, whereas the red curves correspond to the manual ground truth. The initial simple user input is depicted by the blue ellipses in the first line of the figure. The proposed formulation deals successfully with the variations in shape and pose of the discs, although neither a heavy training nor a costly pose optimization are required.

Quantitative performance evaluations: We assessed the similarities between the ground truth and the segmentations obtained with the proposed method over 10 subjects. We used three measures: the Root Mean Squared Error ($RMSE$), the Dice metric (DM), and the correlation coefficient (r). DM is commonly used to measure the similarity (overlap) between the automatically detected and ground-truth regions [2]: $DM = \frac{2\mathbf{A_{am}}}{\mathbf{A_a}+\mathbf{A_m}}$, with $\mathbf{A_a}$, $\mathbf{A_m}$, and $\mathbf{A_{am}}$ corresponding

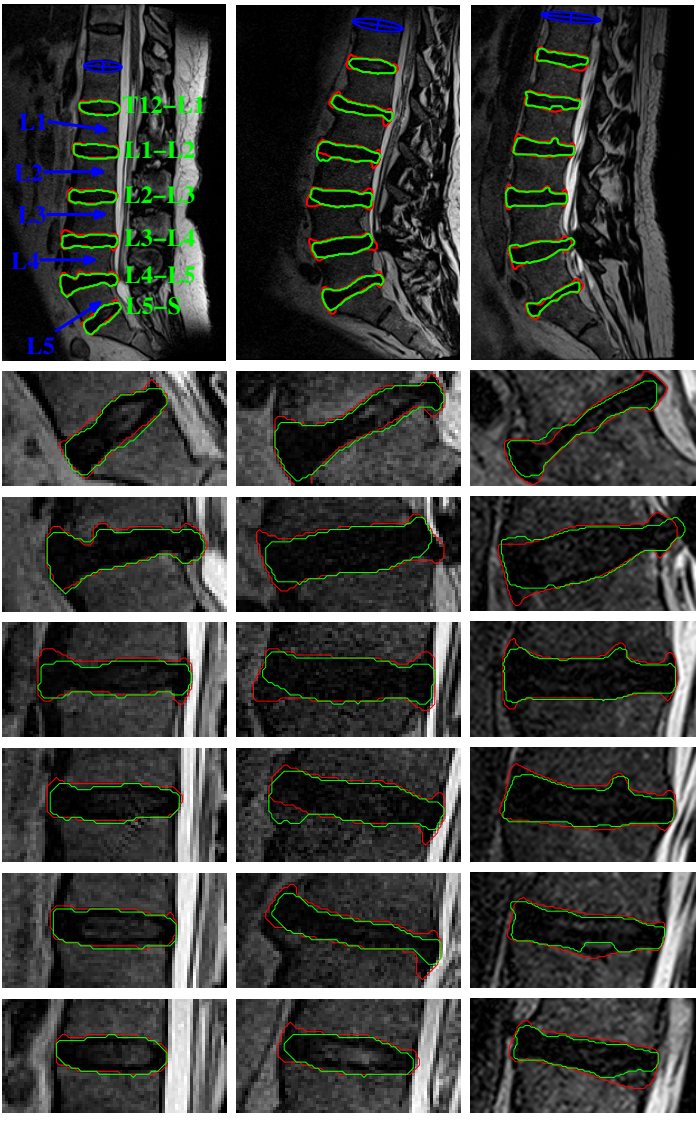


Fig. 3. A representative sample of the results for 3 subjects. Each column depicts the results for one subject: the first line shows the whole image, whereas the rest of the lines show the segmentation results corresponding individual discs from $L5 - S$ to $T12 - L1$ (refer to the standard annotation in the top left image). The green curves depict the segmentations obtained with the proposed algorithm. The red curves correspond to the manual ground truth. The initial simple user input is depicted by the blue ellipses in the first line of the figure. $\lambda = 2.5 \times 10^{-4}$. $\beta = 0.25$.

Table 1. Quantitative performance evaluations over 10 subjects. The parameters were unchanged for all the subjects: $\lambda = 2.5 \times 10^{-4}$ and $\beta = 0.25$.

RMSE mean (in mm)	*DM mean*	*Correlation coefficient (r)*
2.73	0.88	0.98

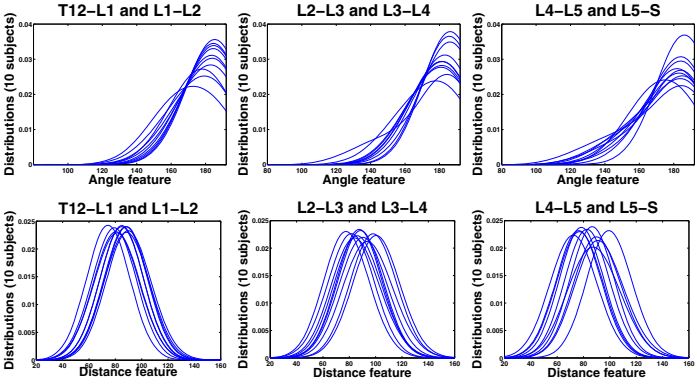


Fig. 4. Invariance of the angle and distance distributions over 10 subjects and 3 different disc pairs (refer to the annotation in the top left image in Fig. 3). The distributions were estimated using 192 bins and a kernel width $\sigma = 15$.

respectively to the areas of the segmented region, the hand-labeled region, and the intersection between them[2]. $RMSE$ evaluates the perpendicular distances from manual to automatic boundaries: $RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \|u_i - v_i\|^2}$, with u_i a point on the automatically detected boundary and v_i the corresponding point on the manually traced boundary (we set $N = 240$)[3]. Table 1 reports the DM mean, the $RMSE$ mean and the correlation coefficient between manual and automatic region areas.

Invariance of the geometric distributions: Using ground truth segmentations, we plotted in Fig. 4 the angle and distance distributions corresponding to 10 different subjects and 3 disc pairs. The figures demonstrate that the angle distributions are very similar (refer to first line in Fig. 4). The distance distributions have similar shapes, but different supports. These shifts, which are due to inter-subject variations in scale, can be handled efficiently with the proposed fixed-point-equation computations. This experiment supports the fact that a single subject is sufficient for training, and confirms the relevance of the proposed priors to disc segmentation.

[2] The higher the DM, the better the performance. DM is always in [0, 1]. $DM > 0.80$ indicates an excellent agreement between manual and automatic segmentations.

[3] The lower $RMSE$, the better the conformity of the results to the ground truth.

References

1. Ben Ayed, I., Chen, H.M., Punithakumar, K., Ross, I., Li, S.: Graph cut segmentation with a global constraint: Recovering region distribution via a bound of the bhattacharyya measure. In: CVPR, pp. 3288–3295 (2010)
2. Ben Ayed, I., Li, S., Ross, I.: Embedding overlap priors in variational left ventricle tracking. IEEE Trans. on Medical Imaging 28(12), 1902–1913 (2009)
3. Boykov, Y., Kolmogorov, V.: An experimental comparison of min-cut/max- flow algorithms for energy minimization in vision. IEEE Trans. on Pattern Analysis and Machine Intelligence 26(9), 1124–1137 (2004)
4. Carballido-Gamio, J., Belongie, S.J., Majumdar, S.: Normalized cuts in 3-d for spinal mri segmentation. IEEE Trans. on Medical Imaging 23(1), 36–43 (2004)
5. Chevrefils, C., Cheriet, F., Aubin, C.E., Grimard, G.: Texture analysis for automatic segmentation of intervertebral disks of scoliotic spines from mr images. IEEE Trans. on Information Technology in Biomedicine 13, 608–620 (2009)
6. Corso, J.J., Alomari, R.S., Chaudhary, V.: Lumbar disc localization and labeling with a probabilistic model on both pixel and object features. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part I. LNCS, vol. 5241, pp. 202–210. Springer, Heidelberg (2008)
7. Huang, S.H., Chu, Y.H., Lai, S.H., Novak, C.L.: Learning-based vertebra detection and iterative normalized-cut segmentation for spinal mri. IEEE Trans. on Medical Imaging 28(10), 1595–1605 (2009)
8. Klinder, T., Ostermann, J., Ehm, M., Franz, A., Kneser, R., Lorenz, C.: Automated model-based vertebra detection, identification, and segmentation in ct images. Medical Image Analysis 13(3), 471–482 (2009)
9. Lee, D.D., Seung, H.S.: Algorithms for non-negative matrix factorization. In: NIPS, pp. 556–562 (2000)
10. Michopoulou, S.K., Costaridou, L., Panagiotopoulos, E., Speller, R., Panayiotakis, G., Todd-Pokropek, A.: Atlas-based segmentation of degenerated lumbar intervertebral discs from mr images of the spine. IEEE Trans. on Biomedical Engineering 56(9), 2225–2231 (2009)
11. Mukherjee, L., Singh, V., Dyer, C.R.: Half-integrality based algorithms for cosegmentation of images. In: CVPR. pp. 2028–2035 (2009)
12. Niemelainen, R., Videman, T., Dhillon, S., Battie, M.: Quantitative measurement of intervertebral disc signal using mri. Clin. Radiol. 63, 252–255 (2008)
13. Rousson, M., Paragios, N.: Prior knowledge, level set representations & visual grouping. International Journal of Computer Vision 76(3), 231–243 (2008)
14. Schmidt, S., Kappes, J.H., Bergtholdt, M., Pekar, V., Dries, S.P.M., Bystrov, D., Schnörr, C.: Spine detection and labeling using a parts-based graphical model. In: Karssemeijer, N., Lelieveldt, B. (eds.) IPMI 2007. LNCS, vol. 4584, pp. 122–133. Springer, Heidelberg (2007)
15. Shi, R., Sun, D., Qiu, Z.L., Weiss, K.: An efficient method for segmentation of mri spine images. In: IEEE ICME, pp. 713–717 (2007)
16. Toshev, A., Taskar, B., Daniilidis, K.: Object detection via boundary structure segmentation. In: CVPR, pp. 950–957 (2010)
17. Vu, N., Manjunath, B.S.: Shape prior segmentation of multiple objects with graph cuts. In: CVPR (2008)