Computer Vision and Image Understanding 117 (2013) 1027–1035

Contents lists available at SciVerse ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

Graph cut segmentation with a statistical shape model in cardiac MRI

D. Grosgeorge [a], C. Petitjean [a,*], J.-N. Dacher [b], S. Ruan [a]

[a] LITIS EA 4108, Université de Rouen, 22 bd Gambetta, 76183 Rouen Cedex, France
[b] INSERM U1096, Université de Rouen, 1 rue de Germont, 76031 Rouen Cedex, France

ARTICLE INFO

Article history:
Received 5 February 2012
Accepted 8 January 2013
Available online 30 April 2013

Keywords:
Image segmentation
Graph cut
Shape prior
MRI
Cardiac ventricle

ABSTRACT

Segmenting the right ventricle (RV) in magnetic resonance (MR) images is required for cardiac function assessment. The segmentation of the RV is a difficult task due to low contrast with surrounding tissues and high shape variability. To overcome these problems, we introduce a segmentation method based on a statistical shape model obtained with a principal component analysis (PCA) on a set of representative shapes of the RV. Shapes are not represented by a set of points, but by distance maps to their contour, relaxing the need for a costly landmark detection and matching process. A shape model is thus obtained by computing a PCA on the shape variations. This prior is registered onto the image via a very simple user interaction and then incorporated into the well-known graph cut framework in order to guide the segmentation. Our semi-automatic segmentation method has been applied on 248 MR images of a publicly available dataset (from MICCAI'12 Right Ventricle Segmentation Challenge). We show that encouraging results can be obtained for this challenging application.

1. Introduction

Magnetic resonance imaging (MRI) is increasingly used as a standard tool in the evaluation of the right ventricle (RV) function [14]. The segmentation of the RV cavity is a prerequisite to the computation of clinical parameters. Although some relatively efficacious methods are commercially available for segmenting the left ventricle (LV) on MR images, such as MASS (Medis, Leiden, The Netherlands) [40] and Argus (Siemens Medical Systems, Germany) [24], the segmentation of the RV is currently performed manually in clinical routine. This lengthy and tedious task requires about 20 min by a clinician and is also prone to intra and inter-expert variability. An automatic segmentation method would avoid these drawbacks, but has to deal with the following issues: (i) fuzziness of the cavity borders due to blood flow and partial volume effect, (ii) the presence of trabeculations (wall irregularities) in the RV, which have the same grey level as the surrounding myocardium, (iii) the complex crescent shape of the RV, which varies according to the imaging slice level (see Fig. 1). This is probably one of the reasons why RV functional assessment has long been considered secondary compared to that of the LV, leaving the problem of RV segmentation wide open [27].

The literature of RV segmentation is thus much less abundant than the one of LV segmentation. Most of RV segmentation methods are based on a joint segmentation of both ventricles. These methods take benefit from the relative positions of the ventricles and the similarity of the gray levels in their respective blood cavities. For example, this information is used within active contours [29], or within a framework combining thresholding, clustering and morphological operations [4]. Another possibility is to use higher-level prior anatomical information to guide the segmentation process, such as a biomechanical model [33] or statistical shape models. In this latter case, the statistical shape model maybe a Point Distribution Model (PDM), obtained by a principal component analysis (PCA) on the set of aligned shapes, and integrated into the well-known active shape and appearance modeling framework [6]. This technique ensures to have a realistic solution since only shapes similar to the training set are allowed, but at the expense of building a training data set with manually generated segmentations. It has been applied to the segmentation of both LV and RV in [23,25]. Statistical prior information may also take the form of an heart atlas. An atlas describes the different structures present in a given type of image. The segmentation of the ventricles is obtained by registering a single [19] or multiple atlases [15] onto the image to be segmented.

Shape prior based segmentation: More generally, automatic organ segmentation can benefit from the use of information regarding shape and/or gray levels, to increase its robustness and accuracy [8]. This especially applies if the nature of the shape does not change much from an example/individual to another. Making use a shape prior for object segmentation involves two steps:

(i) The definition of a mathematical shape representation. In the literature, shapes are usually represented by an explicit

* Corresponding author.
E-mail addresses: damien.grosgeorge@univ-rouen.fr (D. Grosgeorge), caroline.petitjean@univ-rouen.fr (C. Petitjean), su.ruan@univ-rouen.fr (S. Ruan).

1077-3142/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.cviu.2013.01.014

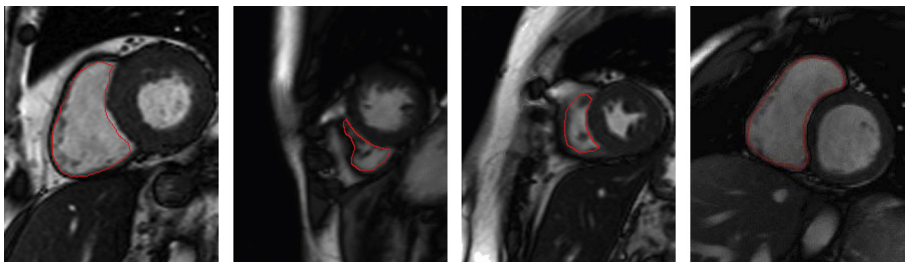

Fig. 1. Typical cardiac MR images. The RV is outlined in red (hand drawn by an expert). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

model such as a point distribution model (PDM) [6] or an implicit one e.g. signed distance function [16]. We discuss both representations, the PDM vs. the SDF in Section 2;

(ii) The definition of the segmentation framework. Initially proposed in the deform models framework, the PDM representation has been mostly used in the active shape and appearance modeling approaches [6]. The SDF representation allowed to easily extend the use of a shape prior into variational frameworks [16,7], and more recently, the graph cut framework [35,11]. In this paper, our aim is to take advantage of the versatility and the low computational cost of the graph cut, in order to propose an efficient prior-based segmentation approach for the RV in cardiac MRI as described in Section 3.

Graph cut with shape prior: The graph cut approach is based on the global optimization of a cost function and is very computationally efficient in 2D [1]. It is also flexible enough to easily take into account shape information. As of today, its use in medical image segmentation has been restricted to a few applications. In [12], a "blob" is created to constraint the segmentation of the entire heart in Computer Tomography (CT) cardiac scans. The shape model may be either an arbitrary fixed template [11], or a parametric curve: an elliptical mask is used for segmenting circular structures (blood vessel) in MRI [35], and circles are used for the LV segmentation in cardiac MRI [46]. But the object to be segmented might not be easily described with a parametric shape. Statistical shape models take into account the variabilities of the shape, at the expense of building a set of manually segmented shapes [36,45]. The use of a specific shape model often imposes an iterative process, with alternates between shape model registration and graph-cut based segmentation, which can be computationally expensive. For example, in [45], the method consists in alternatively searching the PCA, GMM and the pose parameters using gradient descent (maximization step) and segmenting by graph cut using the current shape from the PCA (estimation step), as in an EM framework. Our idea is to design a statistical shape model and to design a graph cut based segmentation framework without the need for an iterative process.

In the remainder of the article, we introduce in Section 2 statistical shape models and two possible representations: point-wise based and signed distance function-based. We propose a shape representation, that will be used for the shape prior. We then recall graph cut basics for image segmentation and introduce our method in Section 3. Experiments provided in Section 4 include a comparison to a state of the art, namely Freedman and Zhang's approach [11]. Conclusion and future works are drawn in Section 5.

2. Representation of statistical shape models

There has been a lot of work on the building and the use of statistical shape models for image segmentation [13,3,31,34,44,17,10]. The use of such models can indeed be of great help when the boundaries of the objet are ill-defined with occlusions. Point-distribution model (PDM) is one of the most widely used representation, used in the ASM framework. The PDM is an explicit shape representation: objects are represented by a finite number of landmarks [5,43,20]. The PDM requires to have point correspondences before performing an analysis. A problem arises in practice when dealing with PDM. The number of available shapes in the training set is often insufficient, where a manual labeling of medical data can be very tedious while establishing correspondences can also be very challenging. This process is prone to variability if manual, and prone to detection error if automatic.

To address this problem, another representation consists in considering the shape as a zero-level set and the values of the remaining voxels by their shortest (usually Euclidian) distance to the boundary [16,7]. This shape representation does not require point correspondences, the shape variability being implicitly represented by the distance variability. Eigenshape decomposition using SDF implicit representation provides tolerance to slight misalignment of object features, since slightly misaligned pixels in a SDF are generally highly correlated. But using a linear (Euclidean) vector space to compute and combine eigenshapes might be an invalid space [30]. However, as point out in [16], the surfaces generally still have advantageous properties of smoothness, local dependence, and zero level sets consistent with the combination of original curves with the use of SDF representation.

In the following section, a statistical shape model based on distance map is proposed, derived from SDF as previously introduced in [16,32]. A comparison with the PDM is carried out to show the similarity of both models.

2.1. Shape representation based on distance function

PDM is an efficient model to represent shape variability. Our idea is to build a statistical shape model which does not need to match corresponding points, but can represent the shape in a similar way as PDM does, as initially proposed in [16]. Consider a set of n two-dimensional binary aligned images of size $H \times W$ composed of shapes of the same class of object. The SDF to each curve is classically defined as $H \times W$ samples encoding the distance to the nearest point on the curve, with for example negative values inside the object. Let $\mathbf{Z}$ be the matrix of SDF of each curve, where each column vector are the $H \times W$ distance samples to the corresponding curve and where each row are the distances for the same location point for each curve of the training set. The objective is to extract shape variations of this matrix $\mathbf{Z}$. The average distance map is considered as the reference by averaging each row of $\mathbf{Z}$:

$$\boldsymbol{\rho} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{Z}_{.\mathbf{i}} \tag{1}$$

$\boldsymbol{\rho}$is a vector of length $H.W$. To capture the variability, we propose a different approach for distance function representation from the one used in the literature [16,37]. Usually, a mean signed distance map is chosen as the reference, which implies that the PCA is performed in the distance function space. As the mean of signed distances is not a signed distance, we reset this map to the distance values. We thus propose to calculate first the mean shape from the mean signed distance maps and then compute its signed distance map. Let p be a point of the image domain included in $\mathbb{R}^2$. The binary mean shape is defined as:

$$\boldsymbol{\chi}_{\rho}(p) = \begin{cases} -1 & \text{if } \boldsymbol{\rho}(p) \geqslant 0 \\ +1 & \text{if } \boldsymbol{\rho}(p) < 0 \end{cases} \tag{2}$$

Let $\mathcal{C}$ be the set of the curve points of the binary mean shape $\boldsymbol{\chi}_{\boldsymbol{\rho}}(p)$. The reference $\boldsymbol{\mu}$ is computed by determining distances to the contour:

$$\boldsymbol{\mu}(p) = \boldsymbol{\chi}_{\boldsymbol{\rho}}(p) \times \inf_{q \in \mathcal{C}} |p - q| \tag{3}$$

Individual shapes are centered:

$$\mathbf{M} = (\mathbf{Z}_{.,1} - \boldsymbol{\mu}) \quad \cdots \quad (\mathbf{Z}_{.,n} - \boldsymbol{\mu}) \tag{4}$$

In this way, PCA is performed in the space of shape variations. The variability in shape is captured using a PCA. Using Singular Value Decomposition (SVD), the covariance matrix, defined by $\frac{1}{n}\mathbf{M}\mathbf{M}^T$, is decomposed to find orthogonal modes of shape variation and corresponding singular values:

$$\frac{1}{n}\mathbf{M}\mathbf{M}^T = \mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^T \tag{5}$$

where $\mathbf{U}$ is a matrix whose column vectors represent the set of orthogonal modes of shape variation, namely eigenvectors, and $\boldsymbol{\Sigma}$ is a $n \times n$ diagonal matrix of corresponding singular values, or eigenvalues. Note that dimension of $\frac{1}{n}\mathbf{M}\mathbf{M}^T$ is large and belong number of pixels, $H.W \times eH.W$. Eigenvectors and eigenvalues computation of this matrix is computationally expensive. A solution exists to deal with large sample size in PCA [38].

In the following, we consider eigenvalues and eigenvectors have been sorted according to size of eigenvalues. Let $k \leqslant n$ be the number of modes to consider, which defines the amount of retained shape variation. Let $\mathbf{z}$ be a novel shape of the same class of object, and $\hat{\mathbf{z}}$ an estimation from $\mathbf{z}$, computed with:

$$\hat{\mathbf{z}} = \boldsymbol{\mu} + \sum_{i=1}^{k} \boldsymbol{\alpha}_i \mathbf{U}_{\mathbf{k}_{.i}} \tag{6}$$

where $\boldsymbol{\alpha}_i$ are obtained with:

$$\boldsymbol{\alpha} = \mathbf{U}_{\mathbf{k}}^{\mathbf{T}}(\mathbf{z} - \boldsymbol{\mu}) \tag{7}$$

with $\mathbf{U}_{\mathbf{k}}^{\mathbf{T}}$ being a matrix consisting of the first k columns of U that is used to project $\mathbf{z}$ into the subspace.

2.2. Shape model comparison

An empirical comparison to illustrate the difference between both representations has been carried out on a dataset of 12 jet images of size 114×114 [37] (Fig. 2). All images are first aligned to the one of them. Both models are then constructed from the aligned images, the PDM being computed on ℓ landmarks. The superposition of both mean shapes and their two principal modes of variation is shown in Fig. 3, that illustrates a similarity of the both mean shapes. Both models do not represent the shape variability in the same space but axes of variation look similar.

In the following, we try to quantify the difference between both models, when used to describe a novel shape instance. Each shape of the jet dataset is being reconstructed with both a SDF and a PDM representation, on a leave-one-out basis: $n - 1$ shapes are used to build both shape models and the last one being used for estimation. Note that the PDM consists of ℓ corresponding landmarks on easily identifiable points. Two standard metrics are then computed to compare the estimated shape and the real one: (i) the Dice coefficient $DM(A, B)$, an overlap measure between two shapes A and B defined by:

$$DM(A, B) = \frac{2|A \cap B|}{|A| + |B|} \tag{8}$$

and (ii) the average point-to-curve ($P2C$) error between two contours, defined by:

$$P2C(A, B) = \frac{1}{|A|}\sum_{a \in A} \min_{b \in B} d(a, b) \tag{9}$$

where $|A|$ denotes the number of points of contour A. Results, shown in Table 1, provide the overlap rates and P2C error in pixels with a varying number of points ℓ and a varying number of eigenvectors k used for the reconstruction. Not surprisingly, reconstruction errors decrease as k increases. Note as well that the PDM seems to capture details slightly better than the SDF representation, whatever the number of points. This can also be observed in Fig. 4, where the reconstruction of two sample shapes by the two representations PDM and SDF are illustrated. Although slightly less accurate, in terms of reconstruction, results obtained by the SDF show that a

Fig. 2. Binary database of 12 aligned jet images (from [37]).

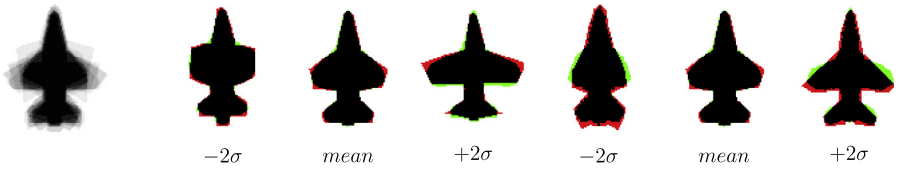

Fig. 3. Shape variability of the fighter jet when using the SDF and the PDM. (a) Overlap image, (b) $\pm 2\sigma$ variations of the first principal mode, (c) $\pm 2\sigma$ variations of the second principal mode. The black part is common to both the SDF and the PDM representation. The red part is part of the PDM only ($\ell = 37$). The green part is part of the SDF only. This figure is best viewed in color. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Mean Dice metric (DM) and P2C (P2C) errors (in pixels) between reconstructed shape and real shape for both the PDM and the SDF representations. DM varies from 0 (total mismatch) to 1 (perfect match). ℓ is the number of points of the PDM. k is the number of modes considered for reconstruction.

		$k = 3$	$k = 6$	$k = 10$
DM	SDF	0.89 ± 0.04	0.91 ± 0.02	0.93 ± 0.02
	PDM $\ell = 16$	0.92 ± 0.03	0.94 ± 0.02	0.96 ± 0.01
	PDM $\ell = 28$	0.92 ± 0.03	0.94 ± 0.02	0.96 ± 0.01
	PDM $\ell = 38$	0.93 ± 0.03	0.94 ± 0.02	0.96 ± 0.01
P2C	SDF	1.76 ± 0.49	1.35 ± 0.33	1.02 ± 0.26
	PDM $\ell = 16$	1.41 ± 0.57	1.06 ± 0.36	0.74 ± 0.23
	PDM $\ell = 28$	1.38 ± 0.56	1.10 ± 0.42	0.71 ± 0.19
	PDM $\ell = 38$	1.34 ± 0.51	1.08 ± 0.40	0.79 ± 0.19

Fig. 4. Two sample shapes reconstructed with a PDM (red curve) and a SDF (green curve). This figure is best viewed in color. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

correct reconstruction of shape can be obtained, without the need for a costly landmark detection and matching process. In summary, the SDF representation has the following advantages over the PDM representation:

- It does not require the positioning of several landmark points, a process prone to variability if manual, and prone to detection error if automatic;
- It does not require landmark matching, a prerequisite before shape analysis, that is difficult to establish;
- It it more robust than the PDM to initial misalignment of the shapes [16].

This SDF representation is thus retained in our segmentation method, described in the next section.

3. Graph cut segmentation with shape prior

In this section, we first outline the graph cut segmentation framework as described in [1]. Then, we introduce the construction of the shape model based on a PCA and show how it can be integrated into the graph cut cost function.

3.1. Graph cut segmentation

Let us consider the image I as a graph $\mathcal{G} = \langle \mathcal{V}, \mathcal{E} \rangle$, where $\mathcal{V}$ is the set of nodes (pixels) and $\mathcal{E}$ the set of edges. Each pair of nodes $(p, q) \in \mathcal{E}$ in a neighborhood $\mathcal{N}$ is connected by a segment called *n-link* and weighted by $B_{p,q}$, a regularization or boundary term, designed to provide spatial coherence in a neighborhood of pixels. $B_{p,q}$ is typically defined as:

$$B_{p,q} \propto \exp\left(-\frac{(I_p - I_q)^2}{2\sigma^2}\right) \cdot \frac{1}{dist(p,q)} \tag{10}$$

where I_p and I_q are the gray levels of pixels p and q, $dist(p, q)$ the Euclidean distance between p and q, and σ a constant usually related to acquisition noise. Consider two additional nodes, called terminal nodes: the source $\mathcal{S}$ representing the object $\mathcal{O}$ (in our case, the RV cavity), and the sink $\mathcal{T}$ representing the background $\mathcal{B}$. Each node $p \in \mathcal{V}$ is connected to the terminal nodes $\mathcal{S}$ and $\mathcal{T}$ by two respective segments called *t-links* and weighted by the so-called region term, denoted R_p and defined by:

$$R_p(\omega) = -\ln Pr(I_p|\omega) \tag{11}$$

where $Pr(I_p|\omega)$ is the likelihood of observing I_p given that pixel p belongs to class ω.

A cut $\mathcal{C}$ in the graph consists in cutting *t-links* and *n-links* to attribute a label $\mathcal{O}$ or $\mathcal{B}$ to each pixel p of the image, which boils down to segmenting the image. The energy of a cut $\mathcal{C}$ is defined by:

$$E(\mathcal{C}) = \sum_{p \in \mathcal{V}} R_p(\omega_p) + \lambda \sum_{p,q \in \mathcal{N}} B_{p,q}.\delta(\omega_p \neq \omega_q) \tag{12}$$

where $\delta(\omega_p \neq \omega_q)$ is 0 if p and q have the same label, 1 otherwise. The optimal segmentation is obtained by searching for the cut of minimal energy. This global search can be very efficiently performed due to mincut-maxflow algorithms, in polynomial time [2].

3.2. Adding the shape prior into the graph cost function

To guide the segmentation process, constraints or models for the object can be introduced through an additional term in the energy formulation of Eq. (12). How to incorporate this prior information depends on the available information: either the constraints are weak and are simple assumptions about the general shape of the object (e.g. convex) or the constraints are strong and concern a precise shape to find in the image.

Weak constraints. In the literature, the constraints on rough shapes are specified through the *n-links*, changing the labeling of the neighboring pixels under the assumption made. In [9], the values of the boundary energy $B_{p,q}$ are modified by prohibiting certain relative positions of p and q, thus promoting compact shapes. The same methodology is used in [41] for more general shapes than convex shapes: it requires that if C is the center of the shape and p a point in the shape, any point q on the line (C, p) after p be also inside the shape. Note that this method has an interesting effect countering shrinkage typically seen in the graph cut segmentation, but implies significant discretization problems. Requiring that the result of segmentation is convex can also be done through an additional energy term as $1 - \cos(\alpha)$, where α is the angle between (p, q) and (p, C) where C is the center of the object designated by a user click [12]. This shows how important angles are penalized by high values of energy, encouraging roughly convex cuts in the graph (Fig. 5).

Strong constraints. When a model of the object is available, it is generally imposed on segmentation through the *t-links*, including an additional term in the energy formulation of the graph which may be similar to the regional term R_p. The classical R_p formulations are changed by replacing the intensity models of the

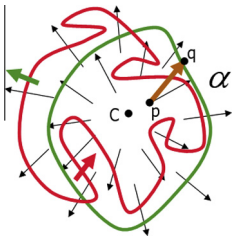

Fig. 5. Example of weak constraint imposed on the graph cut segmentation with calculation of the angle α, allowing the contour (red) to move to a convex shape (green) (from [12]). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

object by labels of the shape prior [35] or a map of probability [36,21], in which case the energy associated with the prior takes the form:

$$E_s(\omega_p) = -\ln(Pr_A(\omega_p))$$

where $Pr_A(\omega_p)$ is the probability of a pixel p belongs to the class ω_p depending on the model. Note that the use of a shape prior implies a difficult problem of matching the model with the image. The registration can be done iteratively and therefore computationaly expensive: the pose parameters estimation and the segmentation calculation are alternated [45,46,42,21]. When registration is a prequisite for the segmentation, process is based on a user interaction [11,36]. The model is defined in this case by a distance map or an atlas. Note that these models are limited to represent judiciously variability in shape.

3.3. Our graph cut segmentation framework using shape prior

A shape model is constructed via a PCA, and summarized in a single map. This map is registered to the image via light user input and is incorporated in the graph cost function. An overview of the method is presented in Fig. 6.

Let us consider N binary shapes of the RV endocardium, obtained by manually segmenting the RV on N cardiac MR images. For each binary shape, a signed distance map ϕ_i to the RV contour is computed. Shapes are rigidly aligned on an arbitrary reference shape and averaged into a mean shape $\bar{\mathbf{\Phi}}$ (Fig. 7a):

$$\bar{\mathbf{\Phi}} = \frac{1}{N}\sum_{i=1}^{N}\phi_{\mathbf{i}} \tag{13}$$

Since averaging does not ensure to obtain a distance function, we propose to reset $\bar{\mathbf{\Phi}}$ to a SDF to the RV contour. A PCA is then performed on the set of centered shapes and yields eigenshapes denoted by $\mathbf{\Phi_i}$, with $i = 1\ldots N$, and their associated eigenvalues, denoted λ_i [37]. A number $k \leqslant N$ of eigenshapes is retained, with k chosen large enough to account for the most important shape variations present in the training set.

Let us now describe how a single prior map is computed from the PCA. Our aim is to isolate areas of variation of the mean shape, for each principal axis. We thus generate highly deformed shape instances for each axis (Fig. 7b and c):

$$\gamma_i^{\pm} = \bar{\mathbf{\Phi}} \pm 3\sqrt{\lambda_i}\mathbf{\Phi_i}, \quad \text{for all } i = 1\ldots k \tag{14}$$

Then, the areas of variation of the mean shape for eigenmode i may be obtained with an exclusive OR between the binary mean shape and the binarized $\gamma_i^{\pm}$:

$$\Gamma_i(p) = \mathcal{H}(\bar{\mathbf{\Phi}}) \oplus \mathcal{H}(\gamma_i^{+}) + \mathcal{H}(\bar{\mathbf{\Phi}}) \oplus \mathcal{H}(\gamma_i^{-}), \quad \text{for all } i = 1\ldots k \tag{15}$$

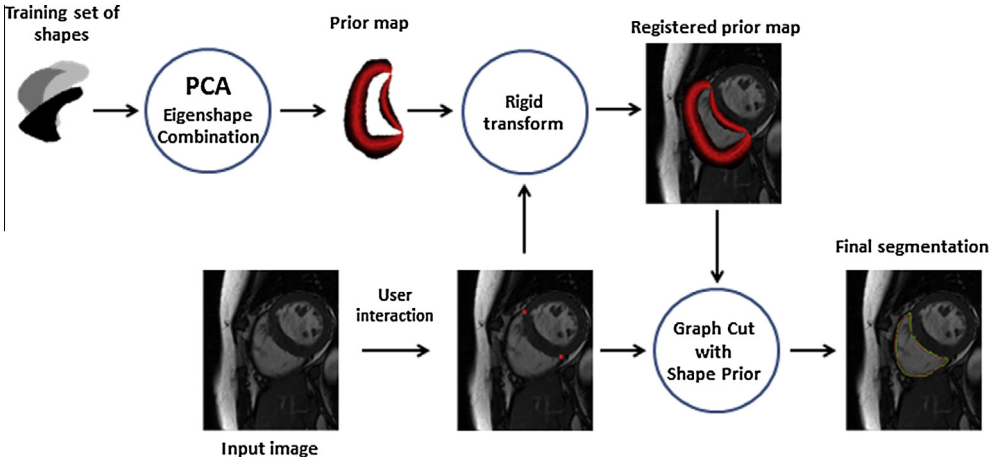

Fig. 6. Overview of the proposed method with shape prior. For the prior map, the darker the higher the distance.

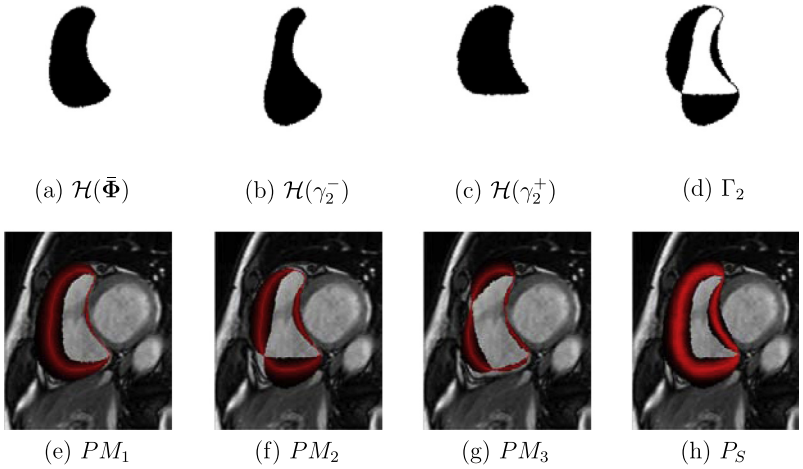

Fig. 7. Computation steps of the prior map. (a) The binary mean shape, (b) and (c) extremal areas of variation of the mean shape for the second axis, (d) final mask for areas of variation for the second axis, (e)–(g) distances to the mean shape for the 3 first axis, (h) final prior map (the darker the higher the distance). This figure is best viewed in color.

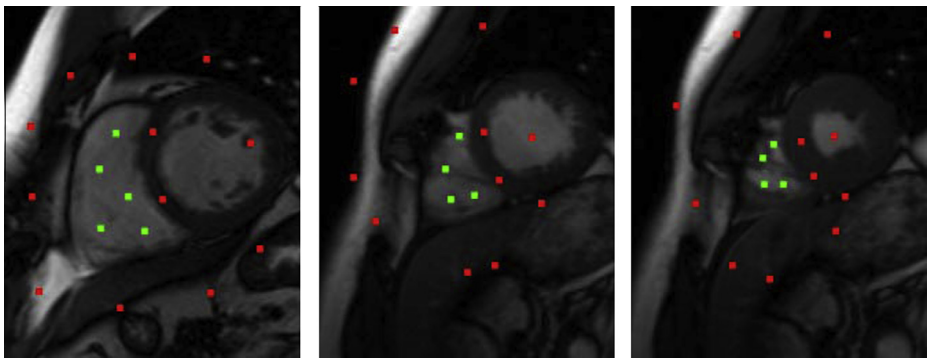

Fig. 8. Typical seeds from the user for object (green) and background (red), used for original graph cut method and Freedman and Zhang's method. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

where $\mathcal{H}(\cdot)$ is the Heaviside function. Γ_i is a binary map that contains areas of variation of the mean shape, for eigenmode i (Fig. 7d). This map is superimposed to the distance values of the mean shape $\bar{\Phi}$ (Fig. 7e and g):

$$PM_i(p) = \Gamma_i(p) \cdot \bar{\Phi}, \quad \text{for all } i = 1 \dots k \tag{16}$$

The k distance maps are averaged into a single distance map (Fig. 7h):

$$P_S(p) = \frac{1}{k}\sum_{i=1}^{k} PM_i(p) \tag{17}$$

P_S includes a distance based region where the contour is expected to be and the complementary region filled with null values.

Now how can this prior map be integrated into the graph cut framework? In the literature, additional energy terms on the *t-links* [36,21] or the *n-links* [11] are added to the graph cost function. In any case, the shape prior must be rigidly registered onto the image to be segmented (see Section 4). We suggest that the shape prior contribute to weighting both *t-links* and *n-links*. The region term R_p can straightforwardly be defined with:

$$R_p^S(\mathcal{O}) = \begin{cases} -\ln(Pr(\mathcal{O}|I_p)) & \text{if } P_S(p) \neq 0 \\ +\infty & \text{if } P_S(p) = 0 \text{ and } \mathcal{H}(\bar{\Phi}(p)) = 1 \text{ (Background)} \\ 0 & \text{if } P_S(p) = 0 \text{ and } \mathcal{H}(\bar{\Phi}(p)) = 0 \text{ (Object)} \end{cases} \tag{18}$$

$$R_p^S(\mathcal{B}) = \begin{cases} -\ln(1 - Pr(\mathcal{O}|I_p)) & \text{if } P_S(p) \neq 0 \\ 0 & \text{if } P_S(p) = 0 \text{ and } \mathcal{H}(\bar{\Phi}(p)) = 1 \text{ (Background)} \\ +\infty & \text{if } P_S(p) = 0 \text{ and } \mathcal{H}(\bar{\Phi}(p)) = 0 \text{ (Object)} \end{cases} \tag{19}$$

with $Pr(\mathcal{O}|I_p)$ a posterior probability model computed from the image gray level pixels p such as $P_S(p) = 0$ and $\mathcal{H}(\bar{\Phi}(p)) = 0$.

Our shape prior is contour-based and may be added as a prior term weighting $n-links$. We thus propose to add a new frontier term denoted by $B_{p,q}^S$ and defined by:

$$B_{p,q}^S = \frac{P_S(p) + P_S(q)}{2} \tag{20}$$

The final energy of a cut $\mathcal{C}$ for a graph integrating a shape prior is then:

$$E(\mathcal{C}) = \lambda \sum_{p,q \in \mathcal{N}} \left(B_{p,q} + \gamma B_{p,q}^S\right).\delta(\omega_p \neq \omega_q) + \sum_{p \in V} R_p^S(\omega_p) \tag{21}$$

where $B_{p,q}$ is defined with Eq. (10), λ weights the relative contributions of the $n-link$ and $t-link$ terms and γ weights the frontier shape prior term $B_{p,q}^S$ and the image frontier term $B_{p,q}$.

4. Experimental results

4.1. Cardiac MR images

Our method has been applied to the datasets of the MICCAI 2012 Cardiac MR Right Ventricle Segmentation Challenge (RVSC).[1] Images have been collected at Rouen University Hospital, as part of the clinical routine. The image database includes 491 short-axis MR images acquired on 32 patients with diagnosed pathologies, who gave written informed consent. For each patient, two volumes a total of 16 images (in average) are available at two time points: 9 images (or slices) at End Diastole (ED, time of maximum filling) and 7 images at End Systole (ES, time of greatest contraction). Training set and Test1 set include both 16 patients. Cardiac images have been zoomed and cropped to a 256×216 (or 216×256) pixel ROI, leaving the LV visible for joint ventricle segmentation, if necessary.

Manual segmentation of the endocardium was performed by a cardiac radiologist, with the convention that trabeculae and papillary muscles were included in the RV cavity. For more information on the data, please refer to [28].

4.2. Shape model construction and method parameterization

The shape models are built based exclusively on the Training Set. There is one model per time point (ED or ES), and several models describing the slice level from base to apex: 6 for ED, 5 for ES. Each of the 11 PCA is performed using between 16 and 32 images. Since size of each eigenvalue indicates the amount of importance its corresponding eigenvector has in determining the shape, we have empirically chosen k to keep 99% of mean shape variations. This corresponds to 7–10 eigenvectors depending on the slice level. Preliminary registration is performed by manually positioning two anatomical landmarks on the interventricular septum (Fig. 6). We have chosen our landmarks based on two criteria, (i) a minimal user interaction, (ii) easily identifiable anatomical landmarks by expert. Based on these criteria, the choice of two points on the junction of the septum to register the model on the 2D images seems to be consistent. Parameters are derived empirically from the training set of 16 patients: $\sigma = 10$ and $\lambda = 100$ for both ED and ES, $\gamma = 0.001$ for ED and $\gamma = 0.005$ for ES. The implementation of Boykov and Kolmogorov of the mincut-maxflow algorithm[2] is used to compute the cut of minimal cost in the graph [2].

[1] http://www.litislab.eu/rvsc.

[2] Available online at http://pub.ist.ac.at/~vnk/software.html.

Table 2
Mean (±standard deviation) Dice Metric (DM) and Point to Curve Distance (P2C, in mm) between automatic and manual delineation of the contour at ED and ES from base (B) to Apex (A) for endocardium.

		Our method		Freedman's method [11]		Original graph cut [1]	
		Dice	P2C (mm)	Dice	P2C (mm)	Dice	P2C (mm)
	B	**0.91 ± 0.09**	**2.25 ± 1.86**	0.87 ± 0.12	4.37 ± 4.12	0.86 ± 0.12	10.99 ± 6.78
		0.90 ± 0.10	**2.31 ± 1.76**	**0.90 ± 0.09**	3.80 ± 3.49	0.88 ± 0.09	10.34 ± 8.12
		0.88 ± 0.12	**2.11 ± 1.80**	0.80 ± 0.18	7.55 ± 7.34	0.77 ± 0.19	15.29 ± 10.61
E		**0.83 ± 0.10**	**2.55 ± 1.28**	0.75 ± 0.19	9.87 ± 9.48	0.71 ± 0.20	20.01 ± 12.40
D		**0.81 ± 0.12**	**2.39 ± 1.39**	0.66 ± 0.23	12.73 ± 11.26	0.61 ± 0.24	23.62 ± 13.54
	A	**0.70 ± 0.18**	**2.27 ± 1.35**	0.56 ± 0.21	12.24 ± 11.14	0.48 ± 0.22	29.62 ± 14.29
	Mean	**0.83 ± 0.15**	**2.32 ± 1.57**	0.74 ± 0.22	8.77 ± 9.37	0.70 ± 0.24	19.22 ± 13.67
	B	**0.84 ± 0.14**	**2.89 ± 2.46**	0.83 ± 0.12	7.12 ± 6.68	0.83 ± 0.08	16.97 ± 9.84
		0.82 ± 0.15	**2.85 ± 1.67**	0.75 ± 0.17	9.06 ± 8.24	0.74 ± 0.21	16.47 ± 9.77
		0.73 ± 0.19	**3.59 ± 1.98**	0.65 ± 0.24	9.67 ± 9.60	0.64 ± 0.23	18.07 ± 11.62
E		**0.66 ± 0.19**	**2.96 ± 1.28**	0.56 ± 0.21	11.62 ± 10.01	0.50 ± 0.21	23.59 ± 13.20
S	A	**0.52 ± 0.21**	**2.86 ± 1.57**	0.39 ± 0.22	19.30 ± 18.79	0.40 ± 0.16	32.30 ± 16.43
	Mean	**0.70 ± 0.22**	**3.05 ± 1.82**	0.61 ± 0.25	12.00 ± 12.98	0.60 ± 0.24	22.26 ± 14.28

Bold values indicate best values for each column.

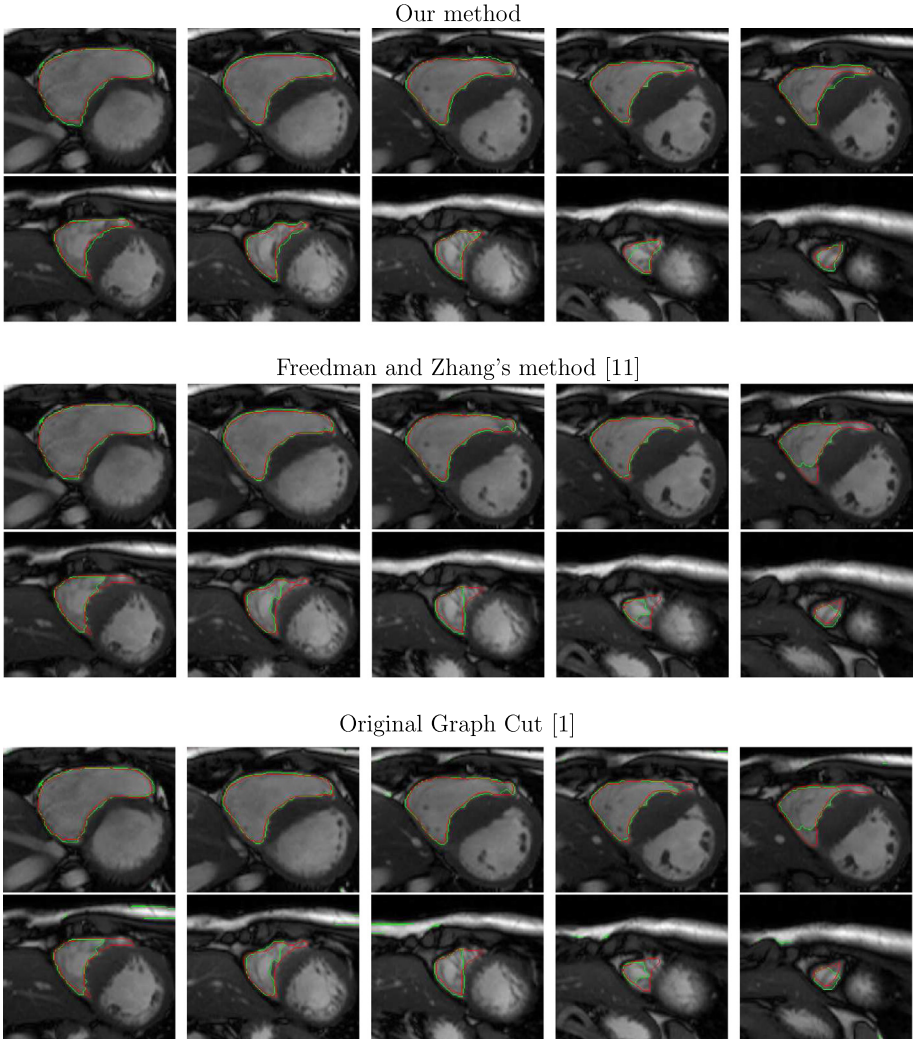

Fig. 9. Segmentation results obtained with automatic algorithm (green) and manual ground truth (red) for Patient 20 of Test1 at ED (from base to apex). Top: our method, middle: Freedman and Zhang's method [11], bottom: original Graph Cut [1]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.3. Segmentation results

Our segmentation algorithm is run on the Test1 Set which contains 16 unseen patients. Our method is compared to one of the first methods for graph cut segmentation with shape prior, as proposed by Freedman and Zhang [11]. Their method consists of using a single shape template described by an unsigned distance map $\bar{\phi}$. We have also compared our method to the original graph cut [1].

Note that both methods require object and background seeds for shape model matching (only in [11]), object and background gray level modeling, and are used as hard constraints in the graph. In average, 5 landmarks have been used for the RV cavity and 10 landmarks for the background (see Fig. 8). Note that parameters have been set empirically from the training set.

For each image of each patient, the user is required to point out two landmarks in order to register the shape prior for our method. Segmentation results for each method are compared to manual ground truth through the Dice Metric (DM) (see Eq. (8)) and the Point to Curve Distance (P2C) (see Eq. (9)). Results are provided in Table 2 and examples of segmentation are shown in Fig. 9.

Accurate segmentation results would be expected to reach intra and inter-observer variability of manual segmentation, that is in the range 1–2 mm [22,26,20,39].[3] Our method yields encouraging results which compare favorably to a state of the art method and outperform the original graph cut approach. Our method provides better results than Freedman and Zhang's, especially in apical images, which are much more difficult to segment (see Fig. 9): they exhibit small structures and are often fuzzy (due to partial volume effect). Freedman and Zhang's prior model might not be specific enough for this kind of images. Regarding basal slices, results are comparable between our method and Freedman and Zhang's, at the cost of substantial user interaction for Freedman and Zhang's method.

Not surprisingly, segmentation results are better for ED images than for ES images, for the three methods: ED images are easier to process, as the heart is then the most dilated. Results are also better for basal and mid-ventricular slices than apical ones. A poor segmentation result in apical slices has little influence on the volume computation but can be a limiting factor in other fields such as studies on the fiber structure. To a lesser extent, this is also true for ES images. Regarding computational cost, our algorithm is implemented in C++ without any particular optimization and requires about 45 s by patient (including both ED and ES volumes) on a Dell E6510 laptop with 4Go RAM and Intel (R) Core (TM) i7 CPU, M460 @ 2.80 GHz. This time is compatible with clinical practice.

5. Conclusion and perspectives

In this paper, we have presented a graph cut based method to segment the RV, using a shape prior. The shape model is constructed via a PCA from a set of representative shapes of the RV obtained by manual segmentation. An original prior term is then integrated into the graphcut cost function. Our segmentation method has been validated on 248 images acquired on 16 patients. It is shown to outperform the native graph cut approach and to compare favorably to the state-of-the-art method. Note however that if results are satisfying for basal and mid-ventricular slices, room for improvement is left in apical slices.

Further investigation concerns a possible 3D extension of our algorithm. Even if our method is scalable to 3D from a theoretical and a memory consuming point of views, a problem arises from the MR clinical data: the space between slices is quite large, typically 8.4 mm, as compared to a spatial resolution of 0.7 mm per pixel in the short-axis view. This raises the general question of the creation of a 3D neighborhood when resolution is anisotropic. One possible solution could be to use a hierarchical approach to decrease time and memory consumption, such as in [18] where a multilevel banded graph cut method is presented. Further investigation will also concern making the registration step automatic.

[3] 1.5 mm and 1.8 mm for endo and epi LV [22], in the range of 1–2 mm on the LV in [26], 1.75 mm for LV and RV in [20] and 1.27 mm and 1.14 mm for LV in [39].

Acknowledgments

The authors thank Dr. Jérôme Caudron for providing the images and for fruitful discussions. They also would the anonymous reviewers whose insightful comments helped to deeply improve the manuscript.

References

[1] Y. Boykov, M. Jolly, Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images, International Conference on Computer Vision 1 (2001) 105–113.
[2] Y. Boykov, V. Kolmogorov, An experimental comparison of min-cut/maxflow algorithms for energy minimization in vision, IEEE Transactions on PAMI 26 (9) (2004) 1124–1137.
[3] G. Charpiat, R. Keriven, O. Faugeras, Shape statistics for image segmentation with priors, in: Conference on Computer Vision and Pattern Recognition, 2007.
[4] C. Cocosco, W.N. Wiro, T. Netsch, E.-J. Vonken, G. Lund, A. Stork, M. Viergever, Automatic image-driven segmentation of the ventricles in cardiac cine MRI, Journal of Magnetic Resonance Imaging 28 (2) (2008) 366–374.
[5] T. Cootes, C. Beeston, G. Edwards, C. Taylor, A unified framework for atlas matching using active appearance models, in: Information Processing in Medical Imaging, LNCS, vol. 1613, 1999, pp. 322–333.
[6] T. Cootes, D. Cooper, C. Taylor, J. Graham, Active shape models – their training and application, Computer Vision and Image Understanding 61 (1) (1995) 38–59.
[7] D. Cremers, T. Kohlberger, C. Schnörr, Nonlinear shape statistics in Mumford-Shah based segmentation, European Conference on Computer Vision (ECCV) (2002) 93–108.
[8] D. Cremers, M. Rousson, R. Deriche, A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape, International Journal of Computer Vision 72 (2) (2007) 195–215.
[9] P. Das, O. Veksler, V. Zavadsky, Y. Boykov, Semiautomatic segmentation with compact shape prior, Image and Vision Computing 27 (2009) 206–219.
[10] R.H. Davies, C.J. Twining, T.F. Cootes, J.C. Waterton, C.J. Taylor, A minimum description length approach to statistical shape modelling, IEEE Transactions on Medical Imaging 21 (5) (2002) 525–537.
[11] D. Freedman, T. Zhang, Interactive graph cut based segmentation with shape priors, in: IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, 2005, pp. 755–762.
[12] G. Funka-Lea, Y. Boykov, C. Florin, M.-P. Jolly, R. Moreau-Gobard, R. Ramaraj, D. Rinck, Automatic heart isolation for CT coronary visualization using graph-cuts, in: IEEE International Symposium Biomedical Imaging (ISBI), 2006, pp. 614–617.
[13] Y. Gao, B. Corn, D. Schifter, A. Tannenbaum, Multiscale 3D shape representation and segmentation with applications to hippocampal/caudate extraction from brain MRI, Medical Image Analysis 16 (2) (2012) 374–385.
[14] F. Haddad, S. Hunt, D. Rosenthal, D. Murphy, Right ventricular function in cardiovascular disease, part i, Circulation 117 (11) (2008) 1436–1448.
[15] H. Kirisli, M. Schaap, S. Klein, L. Neefjes, A. Weustink, T. van Walsum, W. Niessen, Fully automatic cardiac segmentation from 3D CTA data: a multiatlas based approach, in: Proc. SPIE, vol. 7623, 2010, pp. 762305–762309.
[16] M.E. Leventon, W.E.L. Grimson, O. Faugeras, Statistical shape influence in geodesic active contours, in: IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, 2002, pp. 316–323.
[17] W. Liu, Y. Shang, X. Yang, R. Deklerck, J. Cornelis, A shape prior constraint for implicit active contours, Pattern Recognition Letters 32 (15) (2011) 1937–1947.
[18] H. Lombaert, Y. Sun, L. Grady, C. Xu, A multilevel banded graph cuts method for fast image segmentation, in: IEEE International Conference on Computer Vision, vol. 1, 2005, pp. 259–265.
[19] M. Lorenzo-Valdes, G. Sanchez-Ortiz, A. Elkington, R. Mohiaddin, D. Rueckert, Segmentation of 4D cardiac MR images using a probabilistic atlas and the EM algorithm, Medical Image Analysis 8 (3) (2004) 255–265.
[20] J. Lötjönen, S. Kivistö, J. Koikkalainen, D. Smutek, K. Lauerma, Statistical shape model of atria, ventricles and epicardium from short- and long-axis MR images, Medical Image Analysis 8 (3) (2004) 371–386.
[21] J. Malcolm, Y. Rathi, A. Tannenbaum, Graph cut segmentation with nonlinear shape priors, in: IEEE International Conference on Image Processing (ICIP), vol. 4, 2007, pp. 365–368.
[22] L. Marak, J. Cousty, L. Najman, H. Talbot, 4D morphological segmentation and the MICCAI LV-segmentation grand challenge, in: MICCAI 2009 Workshop on Cardiac MR Left Ventricle Segmentation Challenge, MIDAS Journal, 2009.
[23] S. Mitchell, B. Lelieveldt, R. van der Geest, J. Bosch, J. Reiber, M. Sonka, Multistage hybrid active appearance model matching: segmentation of left and right ventricles in cardiac MR images, IEEE Transactions on Medical Imaging 20 (5) (2001) 415–423.
[24] T. O'Donnell, G. Funka-Lea, H. Tek, M.-P. Jolly, M. Rasch, Comprehensive cardiovascular image analysis using MR and CT at siemens corporate research, International Journal of Computer Vision 70 (2) (2006) 165–178.
[25] S. Ordas, L. Boisrobert, M. Huguet, A. Frangi, Active shape models with invariant optimal features (IOF-ASM)-application to cardiac MRI segmentation, in: Computers in cardiology, No. 30, 2003, pp. 633–636.

[26] S. Ordas, H. van Assen, L. Boisrobert, M. Laucelli, J. Puente, B. Lelieveldt, A. Frangi, Statistical modeling and segmentation in cardiac MRI using a grid computing approach, in: Advances in Grid Computing – EGC 2005, Lecture Notes in Computer Science, Springer, Berlin/Heidelberg, 2005, pp. 6–15.
[27] C. Petitjean, J.-N. Dacher, A review of segmentation methods in short axis cardiac MR images, Medical Image Analysis 15 (2) (2011) 169–184.
[28] C. Petitjean, S. Ruan, D. Grosgeorge, J. Caudron, J.-N. Dacher, Right Ventricle Segmentation in Cardiac MRI: a MICCAI'12 Challenge, in: Proc. of 3D Cardiovascular Imaging: a MICCAI segmentation challenge, 2012.
[29] C. Pluempitiwiriyawej, J. Moura, Y. Wu, C. Ho, Cardiac MR image segmentation: quality assessment of STACS, in: IEEE International Symposium on Biomedical Imaging: Macro to Nano. No. 1 in Medical Imaging, April 2004, pp. 828–831.
[30] K. Pohl, J. Fisher, S. Bouix, M. Shenton, R. McCarley, W. Grimson, R. Kikinis, W. Wells, Using the logarithm of odds to define a vector space on probabilistic atlases, Medical Image Analysis 11 (6) (2007) 465–477.
[31] K. Pohl, S. Warfield, R. Kikinis, W. Grimson, W. Wells, Coupling statistical segmentation and PCA shape modeling, in: Medical Image Computing and Computer-Assisted Intervention, LNCS, vol. 3216/2004, 2004, pp. 151–159.
[32] M. Rousson, N. Paragios, Shape priors for level set representations, European Conference on Computer Vision 2 (2002) 78–92.
[33] M. Sermesant, C. Forest, X. Pennec, H. Delingette, N. Ayache, Deformable biomechanical models: application to 4D cardiac image analysis, Medical Image Analysis 7 (4) (2003) 475–488.
[34] T. Shen, H. Li, X. Huang, Active volume models for medical image segmentation, IEEE Transactions on Medical Imaging 30 (3) (2011) 774–791.
[35] G. Slabaugh, G. Unal, Graph cuts segmentation using an elliptical shape prior, in: International Conference on Image Processing (ICIP), vol. 2, 2005, p. 1222.
[36] Z. Song, N. Tustison, B. Avants, J. Gee, Adaptive graph cuts with tissue priors for brain MRI segmentation, in: IEEE International Symposium on Biomedical Imaging (ISBI), 2006, pp. 762–765.
[37] A. Tsai, A. Yezzi, W. Wells, C. Tempany, D. Tucker, A. Fan, W. Grimson, A. Willsky, A shape-based approach to the segmentation of medical imagery using level sets, IEEE Transactions on Medical Imaging 22 (2) (2003) 137–154.
[38] M.A. Turk, A.P. Pentland, Eigenfaces for recognition, Journal of Cognitive Neuroscience 3 (1) (1991) 71–86.
[39] H. van Assen, M. Danilouchkine, A. Frangi, S. Ordas, J. Westenberg, J. Reiber, B. Lelieveldt, SPASM: A 3D-ASM for segmentation of sparse and arbitrarily oriented cardiac MRI data, Medical Image Analysis 10 (2) (2006) 286–303.
[40] R. van der Geest, E. Jansen, V. Buller, J. Reiber, Automated detection of left ventricular epi- and endocardial contours in short-axis MR images, in: Computers in Cardiology. Bethesda, MD, USA, September 1994, pp. 33–36.
[41] O. Veksler, Star shape prior for graph-cut image segmentation, in: European Conference on Computer Vision (ECCV), 2008.
[42] N. Vu, B. Manjunath, Shape prior segmentation of multiple objects with graph cuts, in: IEEE Conference on Computer Vision and Pattern Recognition, 2008.
[43] Y. Wang, L.H. Staib, Boundary finding with correspondence using statistical shape models, in: IEEE Conference on Computer Vision and Pattern Recognition, Santa Barbara, CA, 1998, pp. 338–345.
[44] S. Zhang, Y. Zhan, M. Dewan, J. Huang, D. Metaxas, X. Zhou, Towards robust and effective shape modeling: sparse shape composition, Medical Image Analysis 16 (1) (2012) 265–277.
[45] J. Zhu-Jacquot, R. Zabih, Graph cuts segmentation with statistical shape priors for medical images, in: IEEE Southwest Symposium on Image Analysis and Interpretation, 2008.
[46] J. Zhu-Jacquot, R. Zabih, Segmentation of the left ventricle in cardiac MR images using graph cuts with parametric shape priors, in: International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2008. pp. 521–524.