Max-flow segmentation of the left ventricle by recovering subject-specific distributions via a bound of the Bhattacharyya measure

Ismail Ben Ayed [a,b,*], Hua-mei Chen [b], Kumaradevan Punithakumar [a], Ian Ross [b], Shuo Li [a,b]

[a] GE Healthcare, London, ON, Canada
[b] University of Western Ontario, London, ON, Canada

ABSTRACT

This study investigates fast detection of the left ventricle (LV) endo- and epicardium boundaries in a cardiac magnetic resonance (MR) sequence following the optimization of two original discrete cost functions, each containing global intensity and geometry constraints based on the Bhattacharyya similarity. The cost functions and the corresponding max-flow optimization built upon an original bound of the Bhattacharyya measure yield competitive results in nearly real-time. Within each frame, the algorithm seeks the LV cavity and myocardium regions consistent with subject-specific model distributions learned from the first frame in the sequence. Based on global rather than pixel-wise information, the proposed formulation relaxes the need of a large training set and optimization with respect to geometric transformations. Different from related active contour methods, it does not require a large number of iterative updates of the segmentation and the corresponding computationally onerous kernel density estimates (KDEs). The algorithm requires very few iterations and KDEs to converge. Furthermore, the proposed bound can be used for several other applications and, therefore, can lead to segmentation algorithms which share the flexibility of active contours and computational advantages of max-flow optimization. Quantitative evaluations over 2280 images acquired from 20 subjects demonstrated that the results correlate well with independent manual segmentations by an expert. Moreover, comparisons with a related recent active contour method showed that the proposed framework brings significant improvements in regard to accuracy and computational efficiency.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

The standardized left ventricle (LV) segmentation in Cerqueira et al. (2002) prescribes the use of representative 2D cardiac slices to generate 17 standardized LV segments relevant to the clinical assessment of regional (localized) heart motion abnormalities. Such standard 2D segments are commonly used for regional analysis and quantification of the LV function. An essential component in the diagnosis of cardiovascular diseases related to localized regions with movement abnormalities, obtaining these segments requires accurate detection of the LV endo- and epicardium boundaries in cardiac Magnetic Resonance (MR) sequences (Zhu et al., 2010; Ben Ayed et al., 2009a; Spottiswoode et al., 2009; Hautvast et al., 2006). The problem amounts to segmenting each frame into three target regions: the LV cavity, myocardium, and background (refer to the examples in Fig. 4). Manual delineation in all the images of a subject is prohibitively time-consuming and, as such,

automatic or semi-automatic algorithms are highly desired. Albeit an impressive research effort has been devoted to this problem, current algorithms are still not sufficiently fast and flexible for routine clinical use, mainly because of the challenges inherent to MR cardiac images (Jolly, 2008). For instance, cardiac regions, such as the papillary muscles and the myocardium, are connected and have almost the same intensity profile (cf. the typical example in Fig. 3). Furthermore, appropriate intensity and geometry models of the LV are hard to learn from a finite training set (Ben Ayed et al., 2009a; Hautvast et al., 2006; Jolly, 2008) given the substantial variations in the LV shape and intensity between subjects, particularly those with pathological patterns.

Current algorithms are based, among others, on active contours (Ben Ayed et al., 2009a,c; Hautvast et al., 2006; Fradkin et al., 2008; Lynch et al., 2008; Pluempitiwiriyawej et al., 2005; Sun et al., 2005; El-Berbari et al., 2007; Kaus et al., 2004; Fritscher et al., 2005), active appearance and shape models (Andreopoulos and Tsotsos, 2008; Mitchell et al., 2002; Mitchell et al., 2001; Zambal et al., 2006; Van Assen et al., 2008), classification using probabilistic atlases (Lorenzo-Valdés et al., 2004), graph cuts (Zhu-Jacquot and Zabih, 2008; Ben Ayed et al., 2009), Bayesian fil-

tering (Punithakumar et al., 2010), and registration (Zhuang et al., 2008). Commonly, the problem is stated as the optimization of a cost functional whose solution is obtained following the evolution of an active contour toward the boundary of the target region. Optimization of active contour functionals is a prevalent and flexible choice in medical image analysis because it can introduce a wide spectrum of intensity and geometry constraints[1] on the solution (Ben Ayed et al., 2009b; Rousson and Cremers, 2005; Liu et al., 2005). Generally, these constraints reference a sum over the target region or its boundary of *pixel-wise* penalties fitting the image to intensity and geometry models learned from a finite training set. Pixel-wise intensity information cannot distinguish connected cardiac regions having almost the same intensity profile (Ben Ayed et al., 2009a; Hautvast et al., 2006), for instance the papillary muscles within the cavity and the myocardium. Therefore, most of existing methods bias the solution towards a model of shapes learned *a priori*.

Although very effective in some cases, training-based algorithms may have difficulty in capturing the substantial subject variations in a clinical context (Ben Ayed et al., 2009a; Hautvast et al., 2006; Jolly, 2008). The ensuing results are bounded to the characteristics, variability, and mathematical description of the training set. For instance, a pathological case outside the training set of shapes may not be recovered, and intensity models have to be updated for new acquisition protocols and sequences.

To relax the dependence on the choice of a training set, the recent active curve studies in Zhu et al. (2010), Ben Ayed et al. (2009a,c), Hautvast et al. (2006) build subject-specific models from a user-provided segmentation of one frame in the current cardiac sequence. For instance, in Hautvast et al. (2006), the authors propose to maintain a constant intensity environment in the vicinity of the cavity boundary propagated over the sequence. Based on a global similarity measure between distributions, the method in Ben Ayed et al. (2009a) maintains over a cardiac sequence a constant overlap between the intensity distributions of the cavity and myocardium, which led to promising results for mid-cavity images. In a closely related direction, the authors in Ben Ayed et al. (2009c) maximize the similarities between the intensity distributions of cardiac regions within consecutive frames. Based only on the current data, these methods allow more flexibility in clinical use, although at the price of a user initialization. The LV segmentation methods in Ben Ayed et al. (2009a,c) follow on the effort of several recent studies in the context of general-purpose segmentation (Ben Ayed et al., 2009b, 2010; Ni et al., 2009; Zhang and Freedman, 2005; Freedman and Zhang, 2004; Aubert et al., 2003; Michailovich et al., 2007; Georgiou et al., 2007; Rother et al., 2006), which have shown that the use of *global distribution measures* outperforms standard techniques based on pixel-wise information, and is less sensitive to inaccuracies in estimating the models (Ben Ayed et al., 2009b; Michailovich et al., 2007). As such, it can relax the need of a large training set. Unfortunately, optimization of a global distribution measure with respect to segmentation is *NP-hard* (Rother et al., 2006), and the problem has been commonly addressed with active contour optimization via partial differential equations (Ben Ayed et al., 2009b; Ni et al., 2009; Zhang and Freedman, 2005; Freedman and Zhang, 2004; Aubert et al., 2003; Michailovich et al., 2007; Georgiou et al., 2007). A gradient flow equation of contour evolution is derived in order to increase the similarity between the region within the contour and a given model (Zhang and Freedman, 2005; Freedman and Zhang, 2004; Aubert et al., 2003), or to decrease the similarity between the segmentation

regions defined by the interior and exterior of the contour (Michailovich et al., 2007; Georgiou et al., 2007).

Several measures were studied within the active contour framework, for instance the Kullback–Leibler divergence (Aubert et al., 2003; Zhang and Freedman, 2005; Freedman and Zhang, 2004), the Earth Mover's Distance (Adam et al., 2009), and the Bhattacharyya coefficient (Zhang and Freedman, 2005; Freedman and Zhang, 2004; Ben Ayed et al., 2009b; Michailovich et al., 2007). However, the latter has shown superior performances over other criteria (Zhang and Freedman, 2005; Michailovich et al., 2007). The Bhattacharyya coefficient has a clear geometric interpretation (Comaniciu et al., 2003), outstanding theoretical properties which were well studied in information theory (Aherne et al., 1997), and a fixed (normalized) range which affords a conveniently practical appraisal of the similarity.

These methods based on distribution measures lead to computationally intensive algorithms. Along with an incremental contour evolution, they require a large number of updates of computationally onerous integrals, namely, the distributions of the regions defined by the contour at each iteration and the corresponding measures. Active contour methods rely on stepwise gradient descent. As a result, the ensuing algorithms are notoriously slow, converge to a local minimum, and depend on the choice of an approximating numerical scheme of contour evolution and the corresponding parameters.

In this study, we state the detection of the LV endo- and epicardium boundaries in a cardiac MR sequence as the minimization with respect to a binary variable (labeling) of two original discrete cost functions based on the Bhattacharyya measure, each containing global intensity and geometry constraints. The global constraints measure a similarity between the distributions of the LV cavity and myocardium regions in the current frame and subject-specific models learned from the first frame in the sequence. As we will see in the experiments, the global constraints are more relevant than pixel-wise ones in the context of cardiac regions. Used in conjunction with regularization terms for smooth segmentation boundaries and a hard constraint to ensure the cavity is enclosed within the myocardium, the global constraints yield competitive results over 2280 cardiac images acquired from 20 subjects. However, they do not afford an analytical form amenable to fast 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 solve efficiently the problem, we first propose an original *bound* of the Bhattacharyya measure by introducing an auxiliary labeling. From this bound, we reformulate the problem as the optimization of auxiliary functions by max-flow iterations, thereby obtaining a nearly real-time segmentation algorithm. Then, we demonstrate that the proposed procedure converges. The contributions of this study are not only in the application context but also in the scope of general-purpose segmentation:

1.1. Contributions in the application context

The proposed formulation removes the need of a large training set, handles intrinsically geometric variations of the LV without biasing the solution towards a set of template shapes, relaxes optimization over geometric transformations, and prevents the papillary muscles from being included erroneously in the myocardium.

1.2. Contributions in the general-purpose context

Variables which are global over the segmentation regions have been generally avoided in the context of max-flow optimization (Boykov and Funka Lea, 2006; Rother et al., 2004; Kohli et al.,

---

[1] Geometry constraints reference object shape, position, and size.

2008; Blake et al., 2004). The proposed upper bound affords a potent global description at a very low computational cost. It requires very few max-flow iterations and KDEs to converge, unlike active contours.

Although our cost functions contain constraints that are designed for cardiac image segmentation, the proposed bound optimization can be applied to any cost function in the form of the Bhattacharyya similarity measure. Therefore, it can be used in other applications and can lead to segmentation algorithms which share the flexibility of active contours and computational advantages of max-flow optimization.

The remainder of this paper is organized as follows. The next section details the proposed cost functions and the corresponding bound optimization. In Section 3, we report a quantitative and comparative performance evaluation using several criteria along with the computation time/load, describe a typical example which illustrates explicitly the relevance of global measures for cardiac image segmentation, and give a representative sample of the results for visual inspection. Section 4 contains a conclusion.

2. Formulation

Consider a MR cardiac sequence containing N image functions[2]

$$\mathbf{I}^n(p) = \mathbf{I}^n_p : \mathcal{P} \subset \mathbb{R}^2 \to \mathcal{I} \subset \mathbb{R}, \quad n \in [1 \dots N],$$

with $\mathcal{P}$ the image domain and $\mathcal{I}$ the set of intensity variables. The purpose of this study is to partition the domain $\mathcal{P}$ of each frame n into three regions (cf. the examples in Fig. 4):

$$\mathbf{C}^n : \text{the heart cavity}$$
$$\mathbf{M}^n : \text{the myocardium}$$
$$\mathcal{P} \setminus (\mathbf{C}^n \cup \mathbf{M}^n) : \text{the background}$$

We state the problem as the successive minimization with respect to a binary variable (labeling) of two original *discrete* cost functions designed to address the problems related to cardiac MR images, each containing two *kernel density matching* terms, one intensity-based and the other distance-based. The cost functions are used in conjunctions with regularization terms for smooth partition boundaries and a hard constraint to ensure the cavity region is enclosed within the myocardium. Minimization of the first cost function yields for each frame n an optimal labeling $\mathcal{L}^n_{\mathbf{C}}(p) : \mathcal{P} \to \{0,1\}$ which defines the heart cavity

$$\mathbf{C}^n = \{p \in \mathcal{P}/\mathcal{L}^n_{\mathbf{C}}(p) = 1\}, \tag{1}$$

whereas the labeling minimizing the second cost function, $\mathcal{L}^n_{\mathbf{M} \cup \mathbf{C}}(p) : \mathcal{P} \to \{0,1\}$, defines the myocardium

$$\mathbf{M}^n = \{p \in \mathcal{P}/\mathcal{L}^n_{\mathbf{M} \cup \mathbf{C}}(p) = 1 \quad \text{and} \quad \mathcal{L}^n_{\mathbf{C}}(p) = 0\} \tag{2}$$

2.1. General notations and definitions

We first consider several notations which will be used repeatedly in the definition of the cost functions. For any labeling $\mathcal{L}(p) = \mathcal{L}_p : \mathcal{P} \to \{0,1\}$, any image $\mathbf{J} : \mathcal{P} \to \mathcal{J}$, and any set of variables $\mathcal{J}$, we have:

- $\mathbf{R}^{\mathcal{L}}_1$ and $\mathbf{R}^{\mathcal{L}}_0$ are the complementary regions defined by

$$\mathbf{R}^{\mathcal{L}}_1 = \{p \in \mathcal{P}/\mathcal{L}(p) = 1\}$$
$$\mathbf{R}^{\mathcal{L}}_0 = \{p \in \mathcal{P}/\mathcal{L}(p) = 0\} = \mathcal{P} \setminus \mathbf{R}^{\mathcal{L}}_1 \tag{3}$$

- $\mathbf{P}_{\mathcal{L}\mathbf{J}}$ is the kernel density estimate (KDE) of the distribution of image data $\mathbf{J}$ within region $\mathbf{R}^{\mathcal{L}}_1$:

$$\forall j \in \mathcal{J}, \quad \mathbf{P}_{\mathcal{L}\mathbf{J}}(j) = \frac{\sum_{p \in \mathbf{R}^{\mathcal{L}}_1} K_j(\mathbf{J}_p)}{\mathbf{A}(\mathbf{R}^{\mathcal{L}}_1)}, \tag{4}$$

where $\mathbf{A}(\mathbf{R})$ denotes the number of pixels within a region $\mathbf{R}$ and K_j the Gaussian kernel (σ is the width of the kernel):

$$K_j(\mathbf{J}_p) = \frac{1}{\sqrt{2\pi\sigma^2}} exp^{-\frac{(j-\mathbf{J}_p)^2}{2\sigma^2}}, \tag{5}$$

The Gaussian kernel is commonly used in density estimation (Bishop, 2007). The value of σ controls the smoothness of the distribution estimate. For example, when $\sigma \approx 0$, the Gaussian kernel in (4) approaches the Dirac function, which yields the discrete normalized histogram as estimate. $\sigma > 0$ corresponds to a smooth, continuous distribution estimate (the higher sigma, the smoother the estimate).

- $\mathcal{B}_{\mathcal{J}}(f, g)$ is the *Bhattacharyya* coefficient measuring the amount of overlap (similarity) between two distributions f and g:

$$\mathcal{B}_{\mathcal{J}}(f, g) = \sum_{j \in \mathcal{J}} \sqrt{f(j)g(j)} \tag{6}$$

$\mathcal{B}_{\mathcal{J}}(f, g)$ has a clear geometric interpretation (Comaniciu et al., 2003); it corresponds to the cosine of the angle between the vectors $(f(j), j \in \mathcal{J})^T$ and $(g(j), j \in \mathcal{J})^T$. Thus, it enforces the condition $\sum_{j \in \mathcal{J}} f(j) = 1$ and $\sum_{j \in \mathcal{J}} g(j) = 1$, thereby considering explicitly f and g as distributions. The range of the Bhattacharyya coefficient is $[0; 1]$, 0 corresponding to no overlap between the distributions and 1 to a perfect match. Such fixed (normalized) range affords a conveniently practical appraisal of the similarity. Table 1 summarizes the general notations for any labeling $\mathcal{L}(p) = \mathcal{L}_p : \mathcal{P} \to \{0,1\}$, any image $\mathbf{J} : \mathcal{P} \to \mathcal{J}$, and any set of variables $\mathcal{J}$.

2.2. The cavity-detection cost function

We assume that a detection of the cavity in first frame $\mathbf{I}^1$, i.e., a labeling $\mathcal{L}^1_{\mathbf{c}}$ defining a partition $\{\mathbf{C}^1, \mathcal{P} \setminus \mathbf{C}^1\}$, is given. Using prior information from frame $\mathbf{I}^1$ and labeling $\mathcal{L}^1_{\mathbf{c}}$, the intensity and geometry model distributions of the cavity are learned and embedded in the following distribution matching constraints to segment subsequent frames.

2.2.1. Intensity-matching constraint

Given the learned model distribution of intensity, denoted $\mathcal{M}_{\mathbf{C}\mathcal{I}}$, i.e.,

$$\mathcal{M}_{\mathbf{C}\mathcal{I}} = \mathbf{P}_{\mathcal{L}^1_{\mathbf{c}}, \mathbf{I}^1}, \tag{7}$$

the purpose of this term is to find, for each subsequent frame $\mathbf{I}^n$, $n \in [2 \dots N]$, a region $\mathbf{C}^n$ whose intensity distribution most closely matches $\mathcal{M}_{\mathbf{C}\mathcal{I}}$. To this end, we minimize the following intensity matching function with respect to $\mathcal{L}$:

Table 1
Summary of the general notations for any labeling $\mathcal{L}(p) = \mathcal{L}_p : \mathcal{P} \to \{0,1\}$, any image $\mathbf{J} : \mathcal{P} \to \mathcal{J}$, and any set of variables $\mathcal{J}$.

Notation	Description
$\mathbf{R}^{\mathcal{L}}_1$	Region $\{p \in \mathcal{P}/\mathcal{L}(p) = 1\}$
$\mathbf{R}^{\mathcal{L}}_0$	Region $\{p \in \mathcal{P}/\mathcal{L}(p) = 0\}$
$\mathbf{P}_{\mathcal{L}\mathbf{J}}$	Distribution of $\mathbf{J}$ within $\mathbf{R}^{\mathcal{L}}_1$
$\mathbf{A}(\mathbf{R})$	The number of pixels within $\mathbf{R}$
$\mathcal{B}_{\mathcal{J}}(f, g)$	$\sum_{j \in \mathcal{J}} \sqrt{f(j)g(j)} \in [0; 1]$

---

[2] The number of frames N is typically equal to 20 or 25.

$$\mathcal{L}_{\mathbf{C}}^n = \arg \min_{\mathcal{L}:\mathcal{P}\to\{0,1\}} \mathbf{B}_{\mathbf{C},\mathbf{I}^n}(\mathcal{L}) = -\mathcal{B}_{\mathcal{I}}(\mathbf{P}_{\mathcal{L},\mathbf{I}^n}, \mathcal{M}_{\mathbf{C},\mathcal{I}})$$

$$= -\sum_{i\in\mathcal{I}} \sqrt{\mathbf{P}_{\mathcal{L},\mathbf{I}^n}(i)\mathcal{M}_{\mathbf{C},\mathcal{I}}(i)} \qquad (8)$$

2.2.2. Distance-matching constraint

The purpose of this term is to constrain the segmentation with prior geometric information (shape, scale, and position of the cavity) obtained from the learning frame. Let O_c be the centroid of cavity $\mathbf{C}^1$ in the learning frame and $\mathbf{D}^c(p) = \mathbf{D}_p^c = \frac{\|p - O_c\|}{N_{\mathbf{D}^c}} : \mathcal{P} \to \mathcal{D}$ a *distance image* measuring the normalized distance between p and O_c, with $\mathcal{D}$ the space of distance variables and $N_{\mathbf{D}^c}$ a normalization constant which is computed systematically in order to restrict the values of image $\mathbf{D}^c$ within interval $[0;1]$. Let $\mathcal{M}_{\mathbf{C},\mathcal{D}}$ be the model distribution of distances within the cavity in the learning frame:

$$\mathcal{M}_{\mathbf{C},\mathcal{D}} = \mathbf{P}_{\mathcal{L}_{\mathbf{C}}^1,\mathbf{D}^c} \qquad (9)$$

We propose to find a region $\mathbf{C}^n$ whose distance distribution most closely matches $\mathcal{M}_{\mathbf{C},\mathcal{D}}$ by solving the following minimization problem:

$$\mathcal{L}_{\mathbf{C}}^n = \arg \min_{\mathcal{L}:\mathcal{P}\to\{0,1\}} \mathbf{B}_{\mathbf{C},\mathbf{D}}(\mathcal{L}) = -\mathcal{B}_{\mathcal{D}}(\mathbf{P}_{\mathcal{L},\mathbf{D}^c}, \mathcal{M}_{\mathbf{C},\mathcal{D}})$$

$$= -\sum_{d\in\mathcal{D}} \sqrt{\mathbf{P}_{\mathcal{L},\mathbf{D}^c}(d)\mathcal{M}_{\mathbf{C},\mathcal{D}}(d)} \qquad (10)$$

Note that this geometric prior is invariant to rotation, and embeds *implicitly* uncertainties with respect to scale via the kernel width σ in (5). The higher σ, the more scale variations allowed. In our experiments, $\sigma = 2$ was sufficient to handle effectively variations in the scale of the cavity (cf. the examples in Fig. 4). Based on global rather pixel-wise information, the proposed geometric prior relaxes (1) learning/modeling of geometric characteristics over a large training set and (2) *explicit* optimization with respect to geometric transformations, unlike existing shape priors. It is worth noting that such invariance can be achieved with other descriptions, e.g., invariant moments (Flusser et al., 2009). However, optimizing an invariant moment with respect to segmentation is not amenable to fast solvers. It may result in difficult, computationally expensive optimization problems.

2.2.3. Total cost function

We propose to minimize a cost function containing the intensity and distance matching terms as well as a regularization term for smooth segmentation boundaries. For each $n \in [2...N]$, the algorithm computes the optimal labeling $\mathcal{L}_{\mathbf{C}}^n$ minimizing the following discrete cost function over all $\mathcal{L}: \mathcal{P} \to \{0,1\}$:

$$\mathcal{L}_{\mathbf{C}}^n = \arg \min_{\mathcal{L}:\mathcal{P}\to\{0,1\}} \mathcal{F}_{\mathbf{C},\mathbf{I}^n}(\mathcal{L})$$

$$= \underbrace{\mathbf{B}_{\mathbf{C},\mathbf{I}^n}(\mathcal{L})}_{Intensity\ Matching} + \underbrace{\gamma\mathbf{B}_{\mathbf{C},\mathbf{D}}(\mathcal{L})}_{Geometry\ Matching} + \underbrace{\lambda\mathbf{S}(\mathcal{L})}_{Smoothness} \qquad (11)$$

where $\mathbf{S}(\mathcal{L})$ is a smoothness (regularization) term which ensures label consistency of neighboring pixels (Boykov and Kolmogorov, 2003)

$$\mathbf{S}(\mathcal{L}) = \sum_{\{p,q\}\in\mathcal{N}} \frac{1}{\|p-q\|} \delta_{\mathcal{L}(p)\neq\mathcal{L}(q)} \qquad (12)$$

with

$$\delta_{x\neq y} = \begin{cases} 1 & \text{if } x \neq y \\ 0 & \text{if } x = y \end{cases} \qquad (13)$$

and $\mathcal{N}$ is a 4-neighborhood system containing all unordered pairs $\{p,q\}$ of neighboring elements of $\mathcal{P}$. γ and λ are positive constants balancing the relative contribution of each term.

2.3. Efficient max-flow optimization via an upper bound of the Bhattacharyya measure

The *global* terms $\mathbf{B}_{\mathbf{C},\mathbf{I}^n}(\mathcal{L})$ and $\mathbf{B}_{\mathbf{C},\mathbf{D}}(\mathcal{L})$ in the cost function in Eq. (11) 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 is a challenging and *NP-hard* problem. To optimize efficiently these terms, we first propose an original bound of the Bhattacharyya measure by introducing an auxiliary labeling. From this bound, we reformulate the problem as the optimization of auxiliary functions by max-flow iterations. Then, we formally demonstrate that the proposed procedures converge.

2.3.1. Upper bounds

To introduce our formulation, let us first introduce the following proposition:

Proposition 1. *Given a fixed (auxiliary) labeling $\mathcal{L}^a$, for any labeling $\mathcal{L}$ verifying $\mathbf{R}_1^{\mathcal{L}} \subset \mathbf{R}_1^{\mathcal{L}^a}$, i.e., the foreground region defined by $\mathcal{L}$ is within the foreground region defined by $\mathcal{L}^a$, and $\forall \alpha \in [0,1]$, we have the following upper bound of $\mathbf{B}_{\mathbf{C},\mathbf{I}^n}(\mathcal{L})$*

$$\mathbf{B}_{\mathbf{C},\mathbf{I}^n}(\mathcal{L}) \leqslant \mathcal{J}_{\mathbf{C},\mathbf{I}^n}(\mathcal{L},\mathcal{L}^a,\alpha) = \sum_{p\in\mathbf{R}_0^{\mathcal{L}}} \mathbf{c}_{p,\mathbf{I}^n}(\mathbf{0}) + (1-\alpha)\sum_{p\in\mathbf{R}_1^{\mathcal{L}}} \mathbf{c}_{p,\mathbf{I}^n}(\mathbf{1}), \qquad (14)$$

with $\mathbf{c}_{p,\mathbf{I}^n}(\mathbf{0})$ and $\mathbf{c}_{p,\mathbf{I}^n}(\mathbf{1})$ given for each p in $\mathcal{P}$ by

$$\begin{cases} \mathbf{c}_{p,\mathbf{I}^n}(\mathbf{0}) = \frac{\mathcal{L}_p^a}{\mathbf{A}(\mathbf{R}_1^{\mathcal{L}^a})} \left(\mathbf{B}_{\mathbf{C},\mathbf{I}^n}(\mathcal{L}^a) + \sum_{i\in\mathcal{I}} K_i\left(\mathbf{I}_p^n\right)\sqrt{\frac{\mathcal{M}_{\mathbf{C},\mathcal{I}}(i)}{\mathbf{P}_{\mathcal{L}^a,\mathbf{I}^n}(i)}} \right) \\ \mathbf{c}_{p,\mathbf{I}^n}(\mathbf{1}) = \frac{\mathbf{B}_{\mathbf{C},\mathbf{I}^n}(\mathcal{L}^a)}{\mathbf{A}(\mathbf{R}_1^{\mathcal{L}^a})} \end{cases} \qquad (15)$$

Similarly, we have the following upper bound of $\mathbf{B}_{\mathbf{C},\mathbf{D}}(\mathcal{L})$

$$\mathbf{B}_{\mathbf{C},\mathbf{D}}(\mathcal{L}) \leqslant \mathcal{J}_{\mathbf{C},\mathbf{D}}(\mathcal{L},\mathcal{L}^a,\alpha) = \sum_{p\in\mathbf{R}_0^{\mathcal{L}}} \mathbf{c}_{p,\mathbf{D}}(\mathbf{0}) + (1-\alpha)\sum_{p\in\mathbf{R}_1^{\mathcal{L}}} \mathbf{c}_{p,\mathbf{D}}(\mathbf{1}) \qquad (16)$$

with $\mathbf{c}_{p,\mathbf{D}}(\mathbf{0})$ and $\mathbf{c}_{p,\mathbf{D}}(\mathbf{1})$ given for each p in $\mathcal{P}$ by

$$\begin{cases} \mathbf{c}_{p,\mathbf{D}}(\mathbf{0}) = \frac{\mathcal{L}_p^a}{\mathbf{A}(\mathbf{R}_1^{\mathcal{L}^a})} \left(\mathbf{B}_{\mathbf{C},\mathbf{D}}(\mathcal{L}^a) + \sum_{d\in\mathcal{D}} K_d(\mathbf{D}_p^c)\sqrt{\frac{\mathcal{M}_{\mathbf{C},\mathcal{D}}(d)}{\mathbf{P}_{\mathcal{L}^a,\mathbf{D}^c}(d)}} \right) \\ \mathbf{c}_{p,\mathbf{D}}(\mathbf{1}) = \frac{\mathbf{B}_{\mathbf{C},\mathbf{D}}(\mathcal{L}^a)}{\mathbf{A}(\mathbf{R}_1^{\mathcal{L}^a})} \end{cases} \qquad (17)$$

Proof of proposition 1. We give a proof of Proposition 1 in Appendix A.

Definition 1. $\mathcal{A}(\mathcal{L},\widehat{\mathcal{L}})$ is called auxiliary function of a given cost function $\mathcal{F}(\mathcal{L})$ if it satisfies the following conditions:

$$\mathcal{F}(\mathcal{L}) \leqslant \mathcal{A}(\mathcal{L},\widehat{\mathcal{L}}) \qquad (18)$$
$$\mathcal{A}(\mathcal{L},\mathcal{L}) = \mathcal{F}(\mathcal{L}) \qquad (19)$$

Auxiliary functions are commonly used in the Nonnegative Matrix Factorization (NMF) literature for optimization (Lee and Seung, 2000). Rather than optimizing the cost function, one can optimize iteratively an auxiliary function of the cost function. At each iteration t, this amounts at optimizing over the first variable

$$\mathcal{L}^{(t+1)} = \arg\min_{\mathcal{L}} \mathcal{A}(\mathcal{L},\mathcal{L}^{(t)}) \qquad (20)$$

Thus, by definition of auxiliary function and minimum, we obtain the following monotonically decreasing sequence of the cost function

$$\mathcal{F}(\mathcal{L}^{(t)}) = \mathcal{A}(\mathcal{L}^{(t)},\mathcal{L}^{(t)}) \geqslant \mathcal{A}(\mathcal{L}^{(t+1)},\mathcal{L}^{(t)}) \geqslant \mathcal{F}(\mathcal{L}^{(t+1)}) \qquad (21)$$

Proposition 2. *For $\alpha = 0$, the following function is an auxiliary function of $\mathcal{F}_{\mathbf{C},\mathbf{I}^n}(\mathcal{L})$*

$$\mathcal{A}_{\mathbf{C},\mathbf{I}^n}(\mathcal{L},\mathcal{L}^a,\alpha) = \mathcal{J}_{\mathbf{C},\mathbf{I}^n}(\mathcal{L},\mathcal{L}^a,\alpha) + \gamma\mathcal{J}_{\mathbf{C},\mathbf{D}}(\mathcal{L},\mathcal{L}^a,\alpha) + \lambda\mathbf{S}(\mathcal{L}) \qquad (22)$$

Proof. We give a proof of Proposition 2 in Appendix B. □

Proposition 2 instructs us to consider the following procedure

Minimization of $\mathcal{F}_{\mathbf{C},\mathbf{I}^n}$:

---

begin
- Initialize the auxiliary labeling $\mathcal{L}^a$
- Initialize $\alpha:\alpha = \alpha_0$ with $0 < \alpha_0 < 1$
  repeat
  1. *Optimize the auxiliary function over $\mathcal{L}$*

  $$\mathcal{L}^{(t)} = \arg\min_{\mathcal{L}:\mathbf{R}_1^{\mathcal{L}}\subset\mathbf{R}_1^{\mathcal{L}^a}}\mathcal{A}_{\mathbf{C},\mathbf{I}^n}(\mathcal{L},\mathcal{L}^a,\alpha)$$

  2. *Update $\mathcal{L}^a$ by $\mathcal{L}^a = \mathcal{L}^{(t)}$*
  3. *Decrease $\alpha:\alpha = \alpha^{\rho}$ with $\rho > 1$*
     until *Conversgence;*

The optimal labeling $\mathcal{L}_{\mathbf{C}}^n$ is given by $\mathcal{L}^{(t)}$ at convergence. This optimal labeling defines the cavity in frame n (region $\mathbf{C}^n$) according to Eq. (1).

Convergence proof: When α approaches zero, $\mathcal{A}_{\mathbf{C},\mathbf{I}^n}(\mathcal{L},\mathcal{L}^a,\alpha)$ approaches an auxiliary function of cost function $\mathcal{F}_{\mathbf{C},\mathbf{I}^n}$ and, therefore, the above procedure leads to a monotonically decreasing sequence of $\mathcal{F}_{\mathbf{C},\mathbf{I}^n}$. This comes directly from (21). Since the cost function is lower bounded (because the Bhattacharyya measure is upper bounded by one), the above minimization procedure converges.

2.3.2. Max-flow optimization

Now notice that the auxiliary function $\mathcal{A}_{\mathbf{C},\mathbf{I}^n}$ in *step 1* of the minimization of $\mathcal{F}_{\mathbf{C},\mathbf{I}^n}$ is the sum of *unary* and *pairwise (sub-modular)* penalties. In combinatorial optimization, a global optimum of such sum can be computed efficiently in low-order polynomial time by solving an equivalent max-flow problem (Boykov and Kolmogorov, 2004). Furthermore, the condition that a solution $\mathcal{L}$ should verify $\mathbf{R}_1^{\mathcal{L}} \subset \mathbf{R}_1^{\mathcal{L}^a}$ can be imposed easily by adding a hard constraint (Boykov and Funka Lea, 2006).

To optimize the auxiliary function, it suffices to build a *weighted graph* (an illustration of the graph is depicted in Fig. 1a)

Table 2
Weights of the graph edges for optimizing auxiliary function $\mathcal{A}(\mathcal{L},\mathcal{L}^a,\alpha)$. The last term in w_{p,T_c} is a hard constraint; it is a positive constant K used for each pixel outside $\mathbf{R}_1^{\mathcal{L}^a}$ to ensure the solution $\mathcal{L}$ verifies $\mathbf{R}_1^{\mathcal{L}} \subset \mathbf{R}_1^{\mathcal{L}^a}$. A large K enforces the condition $w_{p,T_c} > w_{p,S_c}$. This means that any pixel outside $\mathbf{R}_1^{\mathcal{L}^a}$ remains connected to terminal T_c (refer to Fig. 1c) and, therefore, remains within the complement of the cavity region.

| Edge	Weight (cost)
w_{p,S_c}	$\mathbf{c}_{p,\mathbf{I}^n}(\mathbf{0}) + \gamma\mathbf{c}_{p,\mathbf{D}}(\mathbf{0})$
w_{p,T_c}	$(1-\alpha)(\mathbf{c}_{p,\mathbf{I}^n}(\mathbf{1}) + \gamma\mathbf{c}_{p,\mathbf{D}}(\mathbf{1})) + K\left(1 - \mathcal{L}_p^a\right)$
$w_{p,q}$	$\frac{\lambda}{\|p-q\|}$

$$\mathcal{G} = \langle\mathbf{N},\mathbf{E}\rangle, \qquad (23)$$

where $\mathbf{N}$ is the set of vertices (nodes) and $\mathbf{E}$ the set of edges connecting these nodes. $\mathbf{N}$ contains a node for each pixel $p \in \mathcal{P}$ and two additional nodes called terminals, one source S_c representing the cavity region $\mathbf{C}^n$ and one sink (T_c) representing the complement of the cavity, $\mathcal{P} \setminus \mathbf{C}^n$. Let $w_{p,q}$ be the weight of the edge connecting neighboring pixels $\{p,q\}$ in $\mathcal{N}$, and w_{p,S_c} and w_{p,T_c} the weights of the edges connecting each pixel p to source S_c and sink T_c, respectively.

The edge weights (costs) in Table 2 correspond to the optimization of $\mathcal{A}_{\mathbf{C},\mathbf{I}^n}$. Choosing these weights and following the max-flow algorithm of Boykov and Kolmogorov (2004), we compute a minimum cut of $\mathcal{G}$, i.e., a subset of edges in $\mathbf{E}$ whose sum of edge weights is minimal and whose removal divides the graph into two disconnected subgraphs, each containing a terminal node (refer to Fig. 1b and c for an illustration). This minimum cut, which assigns each node (pixel) p in $\mathcal{P}$ to one of the two terminals, induces at each iteration t an optimal labeling $\mathcal{L}^t$ which minimizes globally the auxiliary function.

As illustrated in Fig. 1, we are using the Boykov–Kolmogorov max-flow algorithm (Boykov and Kolmogorov, 2004) in the case of 2D grids with a 4-neighborhood system. In this case, it is well known that the algorithm yields a state-of-the-art efficiency (Boykov and Kolmogorov, 2004). However, as noted in Boykov and Kolmogorov (2004), Juan and Boykov (2007), the efficiency of the algorithm decreases when moving from 2D to 3D grids or when using denser (larger neighborhood) grids.

Interpretation of the minimum cut (min-cut/max-flow): Let us write the auxiliary function in terms of the graph weights in Table 2 (we ignore the smoothness and hard constraints for a clear presentation):
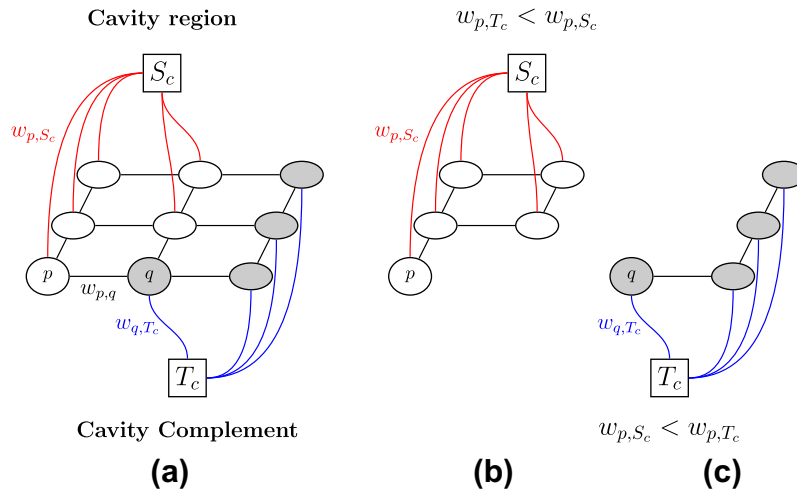
Fig. 1. Illustration of the min-cut/max-flow optimization: the minimum cut yields a partition of the image domain into two regions, one connected to the source terminal and the other to the sink.

$$\mathcal{A}_{\mathbf{C},\mathbf{I}^n} = \sum_{p \in \mathbf{R}_0^{\mathcal{L}}} w_{p,S_c} + \sum_{p \in \mathbf{R}_1^{\mathcal{L}}} w_{p,T_c} \qquad (24)$$

For each pixel p, we have two cases:

Case 1: $w_{p,T_c} < w_{p,S_c}$ (refer to the illustration in Fig. 1b)

In this case, the minimum cut removes the edge with the lower weight w_{p,T_c}, i.e., the edge connecting p and T_c, thereby including pixel p in the cavity region $\mathbf{R}_1^{\mathcal{L}}$ (i.e., foreground). This makes sense because including p in $\mathbf{R}_1^{\mathcal{L}}$ adds the lower weight to the expression in (24) and, therefore, favors the minimization of function $\mathcal{A}_{\mathbf{C},\mathbf{I}^n}$.

Case 2: $w_{p,S_c} < w_{p,T_c}$ (refer to the illustration in Fig. 1c)

In this case, the minimum cut removes the edge connecting p and S_c, thereby excluding pixel p from the cavity region.

A statistical hypothesis testing interpretation of coefficients $\mathbf{c}_{p,\mathbf{I}^n}(\mathbf{0}), \mathbf{c}_{p,\mathbf{I}^n}(\mathbf{1}), \mathbf{c}_{p,\mathbf{D}}(\mathbf{0})\mathbf{c}_{p,\mathbf{D}}(\mathbf{1})$:

In the following, we examine the link between the bound optimization and the classical theory of *hypothesis testing* (Lehmann, 1986). To simplify the interpretation, let us omit the effect of the smoothness, geometry-matching and hard constraints (i.e., $\gamma = \lambda = K = 0$). At the first iteration ($t = 0$), assume that the cavity (foreground) region contains all the pixels in the image. Then, the algorithm performs at each iteration t the following statistical test for each pixel p within the current cavity region ($p \in \mathbf{R}_1^{\mathcal{L}^{t-1}} = \mathbf{R}_1^{\mathcal{L}^a}$):

$$\mathbf{H}_p = w_{p,S_c} - w_{p,T_c} = \mathbf{c}_{p,\mathbf{I}^n}(\mathbf{0}) - (1-\alpha)\mathbf{c}_{p,\mathbf{I}^n}(\mathbf{1})$$

$$= \frac{1}{\mathbf{A}\left(\mathbf{R}_1^{\mathcal{L}^a}\right)}\sqrt{\frac{\mathcal{M}_{\mathbf{C},\mathcal{I}}\left(\mathbf{I}_p^n\right)}{\mathbf{P}_{\mathcal{L}^a,\mathbf{I}^n}\left(\mathbf{I}_p^n\right)}} + \alpha\frac{\mathbf{B}_{\mathbf{C},\mathbf{I}^n}(\mathcal{L}^a)}{\mathbf{A}(\mathbf{R}_1^{\mathcal{L}^a})}$$

$$= \frac{1}{\mathbf{A}\left(\mathbf{R}_1^{\mathcal{L}^a}\right)}\sqrt{\frac{\mathcal{M}_{\mathbf{C},\mathcal{I}}\left(\mathbf{I}_p^n\right)}{\mathbf{P}_{\mathcal{L}^a,\mathbf{I}^n}(\mathbf{I}_p^n)}} - \alpha\frac{\mathcal{B}_{\mathcal{I}}(\mathbf{P}_{\mathcal{L},\mathbf{I}^n},\mathcal{M}_{\mathbf{C},\mathcal{I}})}{\mathbf{A}\left(\mathbf{R}_1^{\mathcal{L}^a}\right)} \qquad (25)$$

The evaluation of the sign of $\mathbf{H}_p$ has a clear meaning, and amounts to a *statistical hypothesis testing* by an image likelihood ratio test. It evaluates the hypotheses that the image at pixel p is drawn from learned model $\mathcal{M}_{\mathbf{C},\mathcal{I}}$ or from the image distribution within the current foreground region. We have two cases:

Case 1: $\mathbf{H}_p < 0$, i.e., the likelihood ratio is lower than the following *critical value*:

$$\frac{\mathcal{M}_{\mathbf{C},\mathcal{I}}\left(\mathbf{I}_p^n\right)}{\mathbf{P}_{\mathcal{L}^a,\mathbf{I}^n}\left(\mathbf{I}_p^n\right)} < (\alpha\mathcal{B}_{\mathcal{I}}(\mathbf{P}_{\mathcal{L},\mathbf{I}^n},\mathcal{M}_{\mathbf{C},\mathcal{I}}))^2 \qquad (26)$$

In this case, we have $\mathbf{c}_{p,\mathbf{I}^n}(\mathbf{0}) < (1-\alpha)\mathbf{c}_{p,\mathbf{I}^n}(\mathbf{1})$, i.e., $w_{p,S_c} < w_{p,T_c}$. Therefore, the graph cut excludes pixel p from the current foreground region (refer to Fig. 1c) so as to decrease the auxiliary function. This makes sense because it results in decreasing the image distribution within the current cavity region at value $\mathbf{I}_p$, which means a better match with the model at that value.

Case 2: $\mathbf{H}_p > 0$

In this case pixel p is kept within the cavity region because excluding it would increase the discrepancy between $\mathbf{P}_{\mathcal{L}^a,\mathbf{I}^n}$ and model $\mathcal{M}_{\mathbf{C},\mathcal{I}}$ at value $\mathbf{I}_p^n$.

Interpretation of α: Parameter α controls the critical value in the right-hand side of (26). The higher α, the more pixels excluded from the current foreground region and, therefore, the faster the evolution of the labeling towards the final segmentation. Therefore α can be viewed as an *algorithmic time* which controls the convergence speed.

2.4. The myocardium-detection cost function

Adopting a similar notation as previously and assuming that a detection of the myocardium in first frame $\mathbf{I}^1$, i.e., a labeling $\mathcal{L}_{\mathbf{M}\cup\mathbf{C}}^1$ defining a partition $\{\mathbf{M}^1 \cup \mathbf{C}^1, \mathcal{P} \setminus \mathbf{M}^1 \cup \mathbf{C}^1\}$, is given, we propose to minimize the following cost function to detect the myocardium:

$$\mathcal{L}_{\mathbf{M}\cup\mathbf{C}}^n = \arg\min_{\mathcal{L}:\mathcal{P}\to\{0,1\}} \mathcal{F}_{\mathbf{M},\mathbf{I}^n}(\mathcal{L})$$

$$= \underbrace{\mathbf{B}_{\mathbf{M},\mathbf{I}^n}(\mathcal{L})}_{\text{Intensity Matching}} + \underbrace{\gamma\mathbf{B}_{\mathbf{M}\cup\mathbf{C},\mathbf{D}}(\mathcal{L})}_{\text{Geometry Matching}} + \underbrace{\lambda\mathbf{S}(\mathcal{L})}_{\text{Smoothness}} + \underbrace{KA\left(\mathbf{R}_0^{\mathcal{L}} \cap \mathbf{R}_1^{\mathcal{L}_{\mathbf{C}}^n}\right)}_{\text{Hard constraint}}$$

$$(27)$$

The first term is an intensity matching term which measures the similarity between $\mathbf{P}_{\mathcal{L},\mathbf{I}^n}$ and a myocardium model of intensity learned from the first frame:

$$\mathbf{B}_{\mathbf{M},\mathbf{I}^n}(\mathcal{L}) = -\mathcal{B}_{\mathcal{I}}(\mathbf{P}_{\mathcal{L},\mathbf{I}^n}, \mathcal{M}_{\mathbf{M},\mathcal{I}}) = -\sum_{i\in\mathcal{I}}\sqrt{\mathbf{P}_{\mathcal{L},\mathbf{I}^n}(i)\mathcal{M}_{\mathbf{M},\mathcal{I}}(i)} \qquad (28)$$

where

$$\mathcal{M}_{\mathbf{M},\mathcal{I}} = \mathbf{P}_{\mathcal{L}_{\mathbf{M}}^1,\mathbf{I}^1} \text{ with } \mathcal{L}_{\mathbf{M}}^1 = \mathcal{L}_{\mathbf{M}\cup\mathbf{C}^1}(1 - \mathcal{L}_{\mathbf{C}}^1) \qquad (29)$$

The second term is a distance matching term:

$$\mathbf{B}_{\mathbf{M}\cup\mathbf{C},\mathbf{D}}(\mathcal{L}) = -\mathcal{B}_{\mathcal{D}}(\mathbf{P}_{\mathcal{L},\mathbf{D}^{mc}}, \mathcal{M}_{\mathbf{MC},\mathcal{D}}) = -\sum_{d\in\mathcal{D}}\sqrt{\mathbf{P}_{\mathcal{L},\mathbf{D}^{mc}}(d)\mathcal{M}_{\mathbf{MC},\mathcal{D}}(d)} \quad (30)$$

where $\mathbf{D}^{mc}(p) = \mathbf{D}_p^{mc} = \frac{\|p-O_{mc}\|}{N_{\mathbf{D}^{mc}}} : \mathcal{P} \to \mathcal{D}$ is a *distance image* measuring the normalized distance between p and O_{mc}, which denotes the centroid of region $\mathbf{M}^1 \cup \mathbf{C}^1$ in the learning frame. $\mathcal{M}_{\mathbf{MC},\mathcal{D}}$ is the model of distances learned from the first frame:

$$\mathcal{M}_{\mathbf{MC},\mathcal{D}} = \mathbf{P}_{\mathcal{L}_{\mathbf{M}\cup\mathbf{C}}^1,\mathbf{D}^{mc}} \qquad (31)$$

The last term is a hard constraint which measures the area of region $\mathbf{R}_0^{\mathcal{L}} \cap \mathbf{R}_1^{\mathcal{L}_{\mathbf{C}}^n}$ (K is a positive constant). Minimization of the area of intersection between $\mathbf{C}^n$ and $\mathbf{R}_0^{\mathcal{L}}$ ensures that the cavity region is enclosed within the myocardium (refer to Fig. 2 for an illustration). The cost function in (27) is minimized by max-flow iterations using previous arguments. The obtained optimal labeling, $\mathcal{L}_{\mathbf{M}\cup\mathbf{C}}^n$, defines the myocardium region in frame n according to Eq. (2).

3. Experimental results and discussions

The evaluation was carried out over 120 short axis cardiac cine MR sequences acquired from 20 subjects: a total of 2280 images were automatically segmented, and the results were compared to
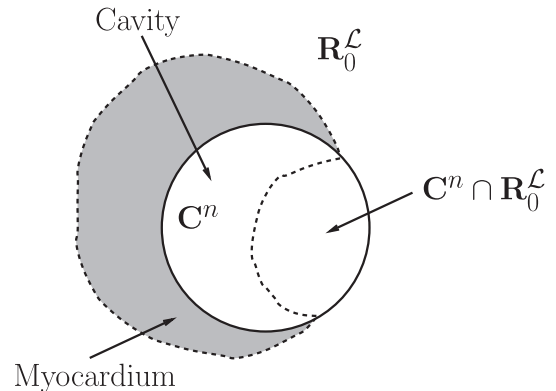
Fig. 2. Illustration of the hard constraint: minimization of the area of intersection between $\mathbf{C}^n$ and $\mathbf{R}_0^{\mathcal{L}}$ ensures that the cavity region is enclosed within the myocardium.

Table 3
Details of the datasets used in evaluation of the proposed method.

Description	Value
Number of subjects	20
Scanner protocol	FIESTA
Patient ages	16–69 years
Short-axis image resolution	(256×256) pixels
Temporal resolution	20 volumes
Pixel spacing	1.17–1.56 mm
Slice thickness	8–10 mm

independent manual segmentations by an expert. For each subject, the algorithm was applied to six slices including apical, mid-cavity, and basal slices. Each slice corresponds to a sequence of 20 frames. Table 3 summarizes the details of the datasets used in evaluation. Using the same datasets, we compared the accuracy and computational load/time of the proposed method, referred to as MFM (max-flow method), with the level-set segmentation of the LV in Ben Ayed et al. (2009a), referred to as LSM (level-set method). Similar to LSM (Ben Ayed et al., 2009a), MFM relaxes the need of a training, and learns model distributions from a user-provided segmentation of the first frame in each sequence. For a fair comparison, the same presegmented learning frame was used for both algorithms.

In the following, we first describe a typical example which illustrates explicitly the relevance of global measures for cardiac image segmentation, and give a representative sample of the results for visual inspection. Then, we report the computation time/load, and describe a quantitative and comparative performance analysis using several accuracy measures. In this analysis, the parameters were unchanged for all the datasets, and were fixed as follows: $\gamma = 1$; $\lambda = 0.0012$; $K = 10^3$; $\sigma = 2$ for the distance distributions, $\sigma = 10$ for the intensity distributions of the cavity, and $\sigma = 10^{-8}$ for the intensity distributions of the myocardium; $\rho = 5$; $\alpha_0 = 0.9$ for the cavity detection and $\alpha_0 = 0.8$ for the myocardium detection. Finally, we will describe a large number of experiments which demonstrate the effect of the weighting parameters (γ and λ) on the results of the proposed algorithm.

3.1. Relevance of global measures: a typical example

Fig. 3 depicts a typical example, where the purpose is to find the boundary between the heart cavity and the myocardium. The expert (ground truth) delineation, depicted by the green curve in Fig. 3b, includes the papillary muscles within the cavity (region

within the green curve). Recovering this expert segmentation is difficult because the myocardium (region outside the green curve) and the papillary muscles within the cavity are connected and have the same intensity profile.

The red curve in Fig. 3b depicts the segmentation obtained by minimizing the pixel-wise likelihood energy commonly used in min-cut segmentation (Boykov and Funka Lea, 2006):

$$\mathcal{L}^{opt} = \arg \min_{\mathcal{L}:\mathcal{P}\to\{0,1\}} -\sum_{p\in\mathbf{R}_0^{\mathcal{L}}} log\mathcal{M}_{\mathbf{M}}(\mathbf{I}_p) - \sum_{p\in\mathbf{R}_1^{\mathcal{L}}} log\mathcal{M}_{\mathbf{C}}(\mathbf{I}_p), \quad (32)$$

where $\mathbf{I}$ is the image data; $\mathcal{M}_{\mathbf{C}}$ and $\mathcal{M}_{\mathbf{M}}$ are the assumed models of image data within the cavity and myocardium, respectively. These models were learned from the ground truth for illustration purposes: $\mathcal{M}_{\mathbf{C}}$ is the distribution of image data within the ground-truth cavity (region inside the green curve) whereas $\mathcal{M}_{\mathbf{M}}$ is the distribution of image data within the ground-truth myocardium (region outside the green curve). Fig. 3c plots these models. The optimization in (32) performs the following test at each pixel p:

Case 1: $-log\mathcal{M}_{\mathbf{M}}(\mathbf{I}_p) < -log\mathcal{M}_{\mathbf{C}}(\mathbf{I}_p)$, i.e., $\mathcal{M}_{\mathbf{M}}(\mathbf{I}_p) > \mathcal{M}_{\mathbf{C}}(\mathbf{I}_p)$

In this case, the pixel is included in the myocardium region (i.e., $\mathbf{R}_0^{\mathcal{L}}$) so as to obtain a lower value of the energy. Therefore, following the optimization in (32), all the cavity pixels corresponding to the gray area in Fig. 3c are erroneously included in the myocardium (refer to the red curve in Fig. 3b).

Case 2: $-log\mathcal{M}_{\mathbf{M}}(\mathbf{I}_p) > -log\mathcal{M}_{\mathbf{C}}(\mathbf{I}_p)$, i.e., $\mathcal{M}_{\mathbf{M}}(\mathbf{I}_p) < \mathcal{M}_{\mathbf{C}}(\mathbf{I}_p)$

In this case, the pixel is included in the cavity region.

Although the models were learned from the ground truth, optimization of the pixel-wise energy in (32) excluded the papillary muscles from the cavity, yielding a segmentation different from the ground truth. Pixel-wise information cannot distinguish between the papillary muscles and the myocardium because the pixels within these two connected regions have almost the same intensity.

Table 4 illustrates explicitly the relevance of the proposed global measure over the pixelwise energy in (32). For each of the segmentations in Fig. 3b, it reports the corresponding pixel-wise energy as well as the global intensity-matching constraint:

$$-\mathcal{B}_{\mathcal{I}}(\mathbf{P}_{\mathcal{L},\mathbf{I}}, \mathcal{M}_{\mathbf{C}}) = -\sum_{i\in\mathcal{I}} \sqrt{\mathbf{P}_{\mathcal{L},\mathbf{I}}(i)\mathcal{M}_{\mathbf{C}}(i)} \quad (33)$$

The ground truth segmentation yields a pixel-wise energy *higher* than the one corresponding to the red-curve segmentation (refer to the second line in Table 4). This indicates that the desired segmentation does not correspond to the minimum of the
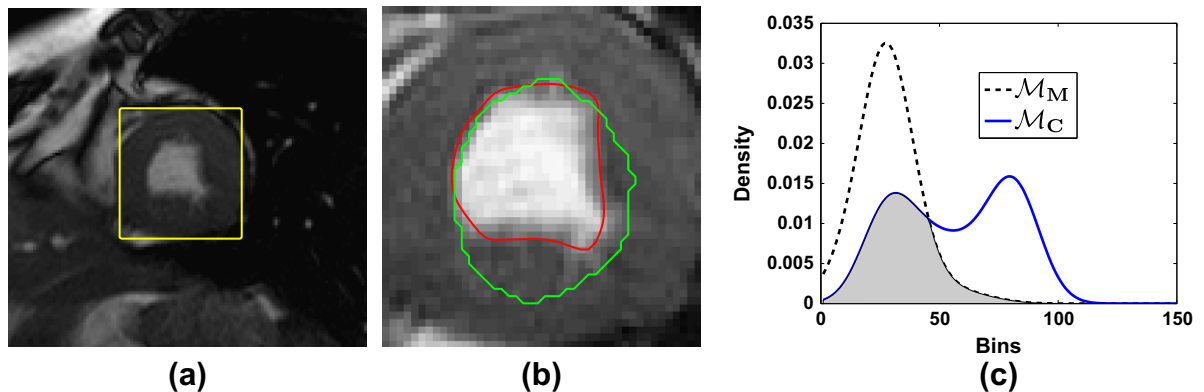
Fig. 3. A typical example where the purpose is to find the boundary between the heart cavity and the myocardium. (a) A box limiting the segmentation to the region of interest. (b) The green curve depicts the expert (ground truth) delineation, and the red curve the segmentation obtained by minimizing the pixel-wise likelihood energy in (32). (c) $\mathcal{M}_{\mathbf{C}}$: the distribution of image data within the ground-truth cavity (region inside the green curve); $\mathcal{M}_{\mathbf{M}}$: the distribution of image data within the ground-truth myocardium (region outside the green curve). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 4
Relevance of global measures: pixel-wise and global energies corresponding to the segmentations in Fig. 3b.

Segmentation	Green curve (ground truth)	Red curve
Global energy	-1	-0.86
Pixel-wise energy	7.38×10^3	7.25×10^3

pixel-wise energy in (32). On the contrary, the global energy (33) corresponding to the ground truth is *lower* than the one corresponding to the red-curve segmentation (refer to the first line in Table 4). This demonstrates explicitly that minimization of the proposed intensity-matching constraint is more relevant for cardiac image segmentation.
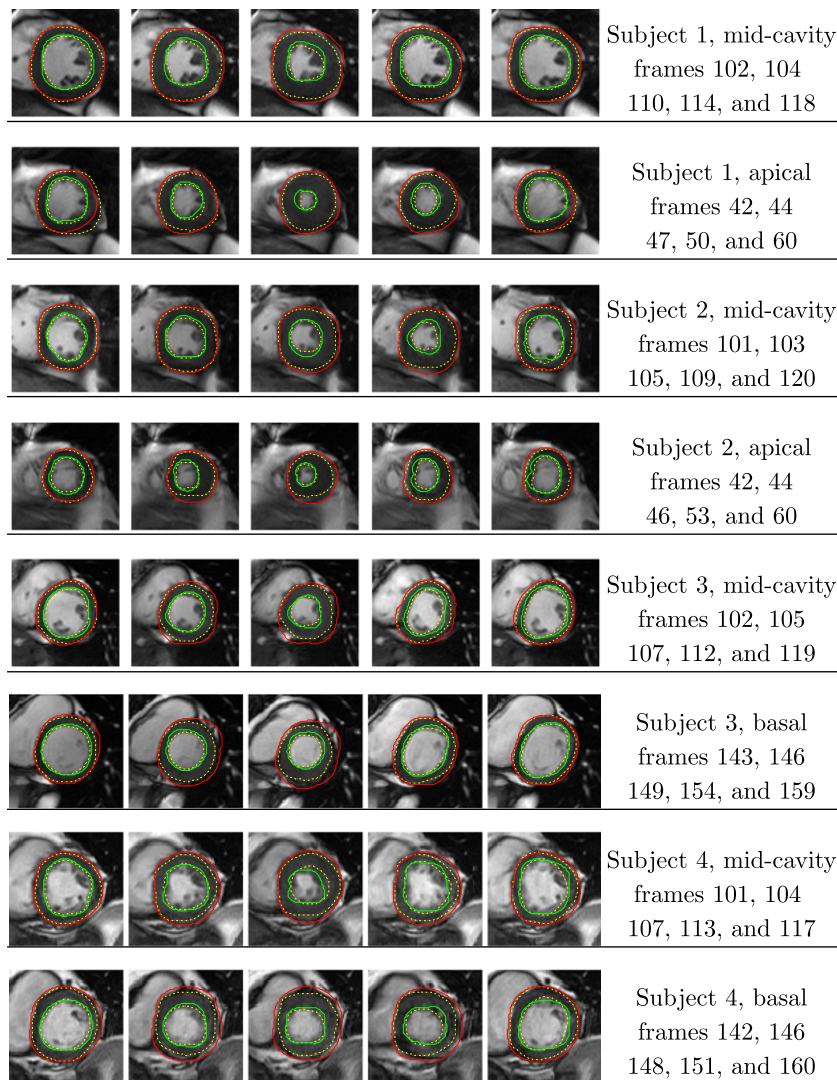
3.2. Visual inspection: a representative sample of the results

In Fig. 4, we give a representative sample of the results for four subjects, including mid-cavity, apical, and basal slices. The green and red curves depict the automatic endo- and epicardium boundaries, respectively. The yellow discontinuous curves depict the ground-truth boundaries. The proposed method prevented the papillary muscles from being included erroneously in the myocardium. This task is challenging (El-Berbari et al., 2007) because the papillary muscles and the myocardium are connected and have almost the same intensity (refer to the green curves in the typical examples depicting mid-cavity frames in Fig. 4). The shown examples also contain apical frames in the second and fourth row in Fig. 4, where it is difficult to segment the cavity because of the small size of the structures and moving artifacts. These examples show how the method handles implicitly significant variations in the scale of the cavity, although neither an additional optimization over geometric transformations nor a large training set are required.

3.3. Computation time/load

Although based on global distribution measures, the proposed method (MFM) led to nearly *real-time* segmentation. Running on a 2 GHz machine, it needs 0.14 s to process a frame. Table 5 reports the computation time/load for NFM and the level-set method in Ben Ayed et al. (2009a) (LSM). The proposed bound leads to a significant decrease in computation load because it requires only two kernel density estimations (KDEs) and max-flow iterations,

	Subject 1, mid-cavity frames 102, 104 110, 114, and 118
	Subject 1, apical frames 42, 44 47, 50, and 60
	Subject 2, mid-cavity frames 101, 103 105, 109, and 120
	Subject 2, apical frames 42, 44 46, 53, and 60
	Subject 3, mid-cavity frames 102, 105 107, 112, and 119
	Subject 3, basal frames 143, 146 149, 154, and 159
	Subject 4, mid-cavity frames 101, 104 107, 113, and 117
	Subject 4, basal frames 142, 146 148, 151, and 160

Fig. 4. A representative sample of the results for four subjects. The yellow discontinuous curves depict the ground-truth boundaries. The green and red curves depict the automatic endo- and epicardium boundaries, respectively. Weighting parameters $\gamma = 1$ and $\lambda = 0.0012$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 5

Computation time (CPU) and number of iterations and kernel density estimations (KDEs) for the proposed method (MFM) and the level-set method in Ben Ayed et al. (2009a) (LSM). MFM led to nearly real-time segmentation and a significant decrease in the computation load.

Method	MFM	LSM
CPU/frame (secs)	0.14	4.33
CPU/subject (secs)	15.96	494.45
Nb of iterations and KDEs/frame	2	300

Table 6

Quantitative performance evaluations over 20 subjects (2280 images) for the proposed method (MFM with $\gamma = 1$ and $\lambda = 0.0012$) and the method in Ben Ayed et al. (2009a) (LSM). The first two rows: The average RMSE. The second two rows: statistics of the DM expressed as mean ± standard deviation (the higher the DM, the better the performance). The last two rows: reliability of the DM ($\mathcal{R}(0.80)$). The higher the reliability, the better the performance.

Method	MFM	LSM
Endocardium RMSE	1.60	2.46
Epicardium RMSE	1.99	1.89
Cavity DM	0.92 ± 0.031	0.88 ± 0.090
Myocardium DM	0.82 ± 0.061	0.81 ± 0.10
Cavity reliability	1	0.89
Myocardium reliability	0.79	0.75

whereas LSM requires approximately 300 KDEs and active-curve updates. KDE allows an accurate and flexible description of model distributions, but is computationally onerous. Generally, the ensuing segmentation algorithms, for instance those based on level-set evolution, compute a large number of updates of kernel densities and, therefore, are computationally intensive. MFM requires very few updates of kernel densities, thereby allowing an accurate and flexible description with a very low computational load.

3.4. Quantitative and comparative performance evaluations

Table 6 summarizes the similarities between the ground truth and the segmentations obtained with MFM ($\gamma = 1$ and $\lambda = 0.0012$) and LSM using three measures: the *Root Mean Squared Error (RMSE)*, the *Dice metric (DM)*, and the correlation coefficient. *RMSE* is contour-based; it measures a distance between manual and automatic boundaries. The Dice metric and correlation coefficient are region-based; they measure the similarity between the automatically detected and ground-truth regions.

3.4.1. The Root Mean Squared Error (RMSE)

We evaluated the *RMSE* by computing the distances between corresponding points on the manual and automatic boundaries. The *RMSE* over *N* points is given by

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\hat{x}_i - \tilde{x}_i)^2 + (\hat{y}_i - \tilde{y}_i)^2} \tag{34}$$

where $(\hat{x}_i, \hat{y}_i)$ is a point on the automatically detected boundary and $(\tilde{x}_i, \tilde{y}_i)$ is the closest point to $(\hat{x}_i, \hat{y}_i)$ on the manually traced boundary. We used 240 points along the boundary, i.e., $N = 240$. *RMSE* measures a distance between manual and automatic boundaries. As such, the lower *RMSE*, the better the conformity of the results to the ground truth. The first two rows in Table 6 report the average *RMSE* in pixels over all the data for MFM and LSM. For endocardium detection, MFM led to a significant improvement in the accuracy over LSM: MFM yielded an average *RMSE* equal to 1.60 pixels, whereas LSM yielded an average *RMSE* equal to 2.46 pixels.

Fig. 5a depicts the average *RMSE* of endocardium detection as a function of the time, i.e., throughout the cardiac cycle, for MFM and

LSM. MFM led to a lower curve, which indicates a better conformity to the ground truth. For epicardium detection, MFM and LSM led approximately to the same accuracy: MFM yielded an average *RMSE* equal to 1.99, whereas LSM yielded an average *RMSE* equal to 1.89. Fig. 5b depicts the average *RMSE* of epicardium detection as a function of the time.

3.4.2. The Dice metric

We computed the *Dice metric* (*DM*) commonly used to measure the similarity (overlap) between the automatically detected and ground-truth regions (Ben Ayed et al., 2009a; Lynch et al., 2008; Pluempitiwiriyawej et al., 2005). The cavity and myocardium regions were evaluated for performance appraisal. Let $\mathbf{V_a}, \mathbf{V_m}$, and $\mathbf{V_{am}}$ be the volumes corresponding to an automatically segmented region, the corresponding hand-labeled region, and the intersection between them, respectively. The volume is measured by the sum of areas of the considered region in six slices. Volume measurements are expressed as the number of pixels within the region. *DM* is given by[3]

$$DM = \frac{2\mathbf{V_{am}}}{\mathbf{V_a} + \mathbf{V_m}} \tag{35}$$

The higher the *DM*, the better the performance of the algorithm.

The second two rows in Table 6 report the *DM* statistics over all the data for MFM and LSM. For cavity detection, MFM led to a significant improvement in region accuracy over LSM: MFM yielded a *DM* equal to 0.92 ± 0.031 (*DM* is expressed as mean ± standard deviation), whereas LSM yielded a *DM* equal to 0.88 ± 0.090. Note that an average *DM* higher than 0.80 indicates an excellent agreement with manual segmentations (Pluempitiwiriyawej et al., 2005), and an average *DM* higher than 0.90 is, generally, difficult to obtain because the small structure of the cavity at the apex decreases significantly the *DM* (Lynch et al., 2008). For myocardium detection, MFM led to a region accuracy slightly better than LSM: MFM yielded a *DM* equal to 0.82 ± 0.061, whereas LSM yielded a *DM* equal to 0.81 ± 0.10.

3.4.3. The reliability

We examined quantitatively and comparatively the *reliability* of the algorithm by evaluating the *reliability function*–i.e., the complementary cumulative distribution function (ccdf)–of the obtained Dice metrics, defined for each $d \in [0, 1]$ as the probability of obtaining *DM* higher than *d* over all volumes:

$$\mathcal{R}(d) = Pr(DM > d)$$
$$= \frac{Number\ of\ volumes\ segmented\ with\ DM > d}{Total\ number\ of\ volumes} \tag{36}$$

$\mathcal{R}(d)$ measures how reliable the algorithm in yielding accuracy *d*, i.e., a *DM* higher than *d*. The higher $\mathcal{R}$, the better the performance. The last two rows in Table 6 reports $\mathcal{R}(0.80)$ for MFM and LSM, indicating *the proposed algorithm brings* 11% *improvement in the reliability of cavity detection and* 4% *improvement in the reliability of myocardium detection*. For cavity detection, MFM yielded $\mathcal{R}(0.80) = 1$, i.e., an excellent agreement (*DM* > 0.80) in 100% of the cases, whereas LSM achieved 89% of the cases with a similar accuracy. For myocardium detection, MFM yielded $\mathcal{R}(0.80) = 0.79$, whereas LSM yielded $\mathcal{R}(0.80) = 0.75$.

In Fig. 6a and b, we plotted $\mathcal{R}$ as a function of *d* for the cavity and myocardium, respectively. The proposed algorithm led to reliability curves higher than LSM. Fig. 6c and d depict the *DM* for a representative sample of the analyzed volumes for MFM and LSM.

---

[3] *DM* is always in [0,1]. *DM* equal to 1 indicates a perfect match between manual and automatic segmentations.
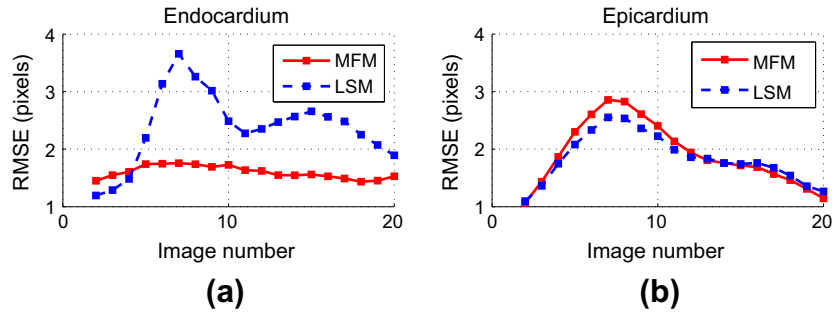
Fig. 5. Average *RMSE* (in pixels) over 20 subjects (2280 images) as a function of the time step for the proposed method (MFM) and the method in Ben Ayed et al. (2009a) (MFM): (a) endocardium detection; (b) epicardium detection. The average frame number corresponding to end diastole is approximately equal to 7.
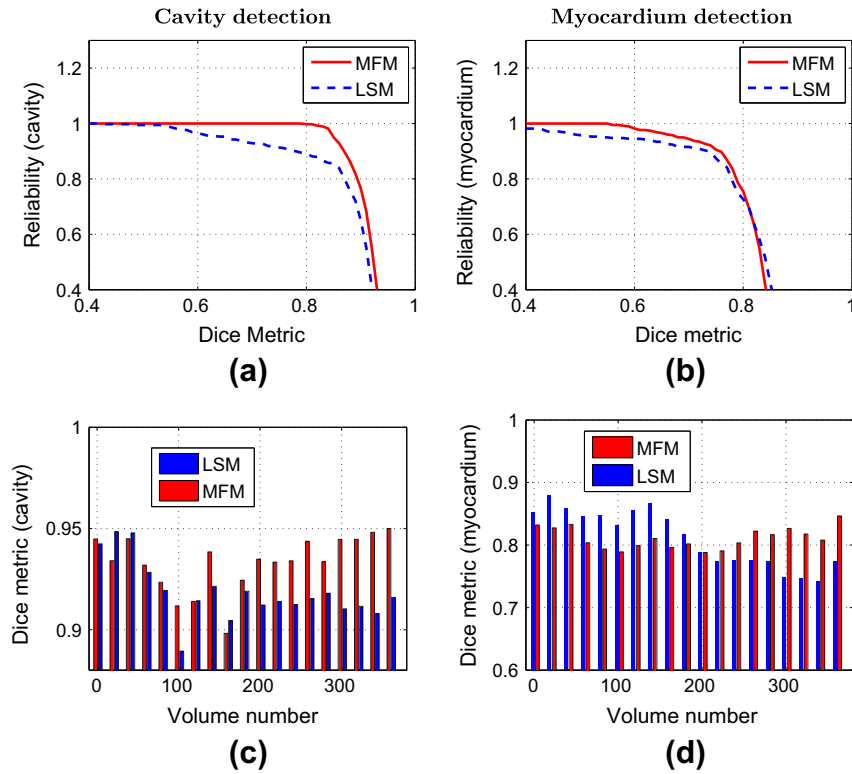
Fig. 6. Region-based comparisons of manual and automatic segmentations of 2280 images (380 volumes) acquired from 20 subjects. (a) and (b): Reliability ($\mathcal{R}(d) = Pr(DM > d)$) for the proposed method (MFM) and the level set method in Ben Ayed et al. (2009a) (LSM). MFM led to reliability curves higher than LSM. (c) and (d): *DM* in a representative sample of the tested volumes for MFM and LSM. The first column corresponds to the cavity detection, and the second to the myocardium detection.
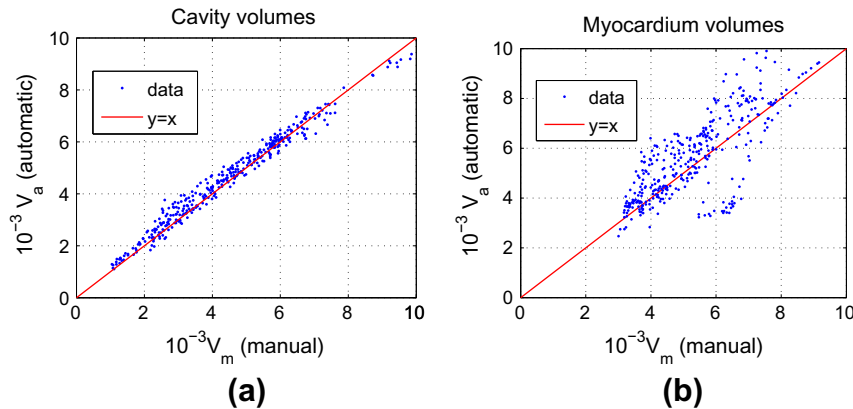
Fig. 7. Automatic volumes versus manual volumes for the proposed method (MFM). (a) Cavity detection: the obtained correlation coefficient is 0.99. (b) Myocardium detection: the obtained correlation coefficient is 0.81.

(a) $\gamma = 0$ (b) $\gamma = 1$
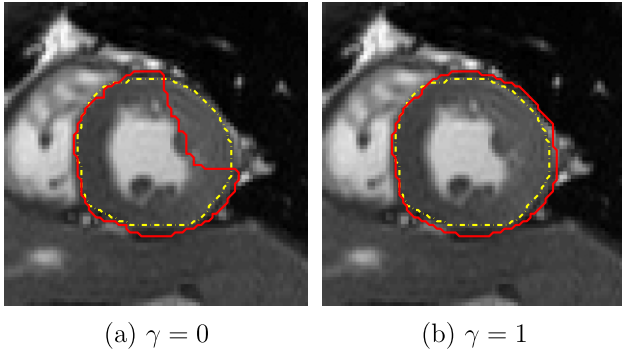
Fig. 8. Effect of the geometry-matching term: The yellow discontinuous curves depict the ground truth boundary; (a) The red curve depicts the epicardium boundary obtained without geometry-matching term ($\gamma = 0$); (b) The red curve depicts the epicardium boundary obtained with the geometry-matching term ($\gamma = 1$); The smoothness weight is fixed for the two experiments ($\lambda = 0.00075$). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 7

Quantitative performance evaluations over 20 subjects (2280 images) for MFM without geometry-matching term ($\gamma = 0$) and MFM without smoothness term ($\lambda = 0$). The first two rows: The average *RMSE*. The second two rows: statistics of the Dice metric (*DM*) expressed as mean ± standard deviation (the higher the *DM*, the better the performance). The last two rows: reliability of the *DM* ($\mathcal{R}(0.80)$). The higher the reliability, the better the performance.

| Method	MFM ($\gamma = 0$)	MFM ($\lambda = 0$)
Endocardium *RMSE*	2.36	3.69
Epicardium *RMSE*	2.89	2.02
Cavity *DM*	0.76 ± 0.2	0.88 ± 0.07
Myocardium *DM*	0.77 ± 0.12	0.80 ± 0.06
Cavity reliability	0.61	0.89
Myocardium reliability	0.52	0.61

3.4.4. The correlation coefficient

The proposed method yielded high correlation coefficients between manual and automatic volumes. The correlation coefficient is equal to 0.99 for cavity detection, and is equal to 0.81 for myocardium detection. The linear regression plots in Fig. 7 illustrate this high correlation. For all the data analyzed (380 volumes), it depicts the volumes obtained with the proposed method versus manual volumes, along with the identity line. Most of the data points are very close to the identity line, which illustrates small differences between manual and automatic segmentations.

### 3.5. Effect of the choice of the weighting parameters

In this section, we examine the role of the different terms in the cost function, and evaluate quantitatively the robustness of the proposed algorithm with respect to the choice of the weighting parameters. First, we start with a typical example which demonstrates explicitly the positive effect of the geometry-matching term. Then, we examine the performance of the algorithm without smoothness ($\lambda = 0$) or without geometry-matching ($\gamma = 0$). Finally, we describe extensive experiments which evaluate the performance of the algorithm as a function of the parameters (Figs. 9 and 10).

The example in Fig. 8 demonstrates explicitly the effect of the geometry-matching term. The yellow discontinuous curve depicts the ground truth epicardium boundary. The red curve in (a) depicts the epicardium boundary obtained without geometry-matching term ($\gamma = 0$). In this case, a part of myocardium region was excluded erroneously from the solution. On the contrary, adding the geometry-matching term ($\gamma = 1$) biased the solution towards the ground-truth (refer to Fig. 8b).

Table 7 reports the performance of the proposed algorithm with $\gamma = 0$. Removing the geometry-matching constraint affected significantly all the performance measures. Table 7 also reports the performance of the proposed algorithm without smoothness
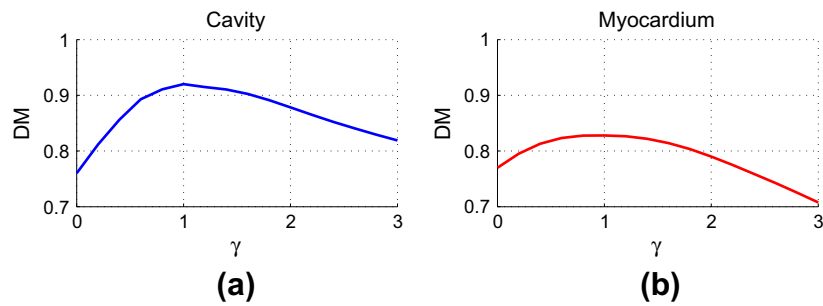
Fig. 9. The average Dice metric over all the data as a function of γ (the weight of the geometry-matching constraint).
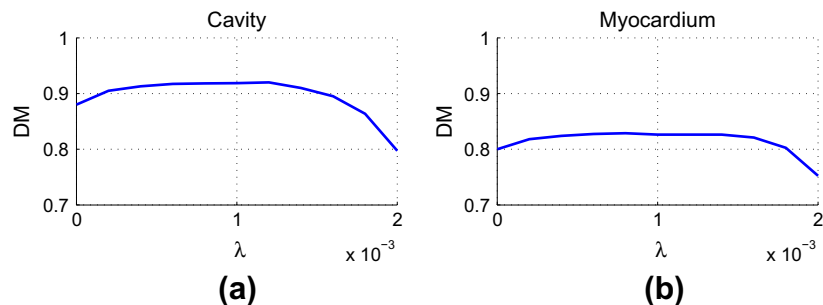
Fig. 10. The average Dice metric over all the data as a function of λ (the weight of the smoothness constraint).

constraint ($\lambda = 0$). Removing the smoothness constraint did not change significantly the average *DM*. However, it affected the *RMSE* and *Reliability*.

Fig. 9a and b plot the average Dice metric as a function of γ (the weight of the geometry-matching constraint) for the cavity and myocardium, respectively. We run the algorithm over 16 uniformly-spaced values of γ in the interval [0; 3]. The values of γ in [0.5; 1.5] yielded a Dice metric above 0.9. Therefore, this parameter does not require a fine tuning. It is interesting to notice that the best performance corresponds to $\gamma = 1$. This makes sense because the intensity and geometry-matching constraints have the same form of a Bhattacharyya measure and, therefore, the same range of values.

Fig. 10a and b plot the average Dice metric as a function of λ (the weight of the smoothness constraint) for the cavity and myocardium, respectively. We run the algorithm over 11 uniformly-spaced values of $\lambda \times 10^3$ in the interval [0; 2]. Most of the values of λ within this interval yield a Dice metric above 0.9 for the cavity and above 0.8 for the myocardium. The Dice metric starts to drop significantly only at the neighborhood of $\lambda \times 10^3 = 2$. Fig. 10a and b demonstrate that the smoothness parameter does not require fine tuning.

4. Conclusion

This study investigated nearly real-time detection of the LV endo- and epicardium boundaries in a cardiac magnetic resonance (MR) sequence. The solution is obtained following the optimization of two original discrete cost functions, each containing global geometry and intensity constraints based on the Bhattacharyya similarity. Quantitative evaluations over 2280 images acquired from 20 subjects demonstrated that the results correlate well with independent manual segmentations by an expert. Compared to the recent active contour method in Ben Ayed et al. (2009a), the proposed formulation led to improvements in accuracy and a significant decrease in computation time. Built upon an original bound of the Bhattacharyya measure, the proposed formulation afford an important computational advantage over related active contour methods: it does not require a large number of iterative updates of the segmentation and the corresponding kernel densities. Another interesting aspect of the bound-optimization framework is that it can be used for several other applications and, therefore, can lead to segmentation algorithms which share the flexibility of active contours and computational advantages of max-flow optimization. Apart from these advantages over related general-purpose segmentation methods, the proposed formulation has several desirable properties in the application context. Based on global distribution information learned from the current data, it removes the need of a large training set, handles intrinsically geometric variations of the LV without biasing the solution towards a set of template shapes, relaxes optimization over geometric transformations, and prevents the papillary muscles from being included erroneously in the myocardium.

Acknowledgments

Appendix A. Proof of Proposition 1

Because $\mathbf{R}_1^{\mathcal{L}}$ and $\mathbf{R}_0^{\mathcal{L}}$ are complementary, we can rewrite $\mathbf{R}_1^{\mathcal{L}^a}$ as follows:

$$\mathbf{R}_1^{\mathcal{L}^a} = \left(\mathbf{R}_1^{\mathcal{L}^a} \cap \mathbf{R}_1^{\mathcal{L}}\right) \cup \left(\mathbf{R}_1^{\mathcal{L}^a} \cap \mathbf{R}_0^{\mathcal{L}}\right) \tag{A.1}$$

For $\mathcal{L}$ verifying $\mathbf{R}_1^{\mathcal{L}} \subset \mathbf{R}_1^{\mathcal{L}^a}$, we have $\mathbf{R}_1^{\mathcal{L}^a} \cap \mathbf{R}_1^{\mathcal{L}} = \mathbf{R}_1^{\mathcal{L}}$. Therefore, from Eq. (A.1), we can rewrite $\mathbf{R}_1^{\mathcal{L}}$ as follows:

$$\mathbf{R}_1^{\mathcal{L}} = \mathbf{R}_1^{\mathcal{L}^a} \setminus \left(\mathbf{R}_1^{\mathcal{L}^a} \cap \mathbf{R}_0^{\mathcal{L}}\right) \tag{A.2}$$

Using this equation, we rewrite the kernel density estimate in (4) as follows:

$$\mathbf{P}_{\mathcal{L},\mathbf{I}^n}(i) = \frac{\sum_{p \in \mathbf{R}_1^{\mathcal{L}^a}} K_i\left(\mathbf{I}_p^n\right) - \sum_{p \in \mathbf{R}_1^{\mathcal{L}^a} \cap \mathbf{R}_0^{\mathcal{L}}} K_i\left(\mathbf{I}_p^n\right)}{\mathbf{A}\left(\mathbf{R}_1^{\mathcal{L}^a}\right) - \mathbf{A}\left(\mathbf{R}_1^{\mathcal{L}^a} \cap \mathbf{R}_0^{\mathcal{L}}\right)} \tag{A.3}$$

Now because $\mathbf{A}\left(\mathbf{R}_1^{\mathcal{L}^a} \cap \mathbf{R}_0^{\mathcal{L}}\right)$ is nonnegative, we have the following inequality

$$\mathbf{P}_{\mathcal{L},\mathbf{I}^n}(i) \geqslant \frac{\sum_{p \in \mathbf{R}_1^{\mathcal{L}^a}} K_i\left(\mathbf{I}_p^n\right) - \sum_{p \in \mathbf{R}_1^{\mathcal{L}^a} \cap \mathbf{R}_0^{\mathcal{L}}} K_i\left(\mathbf{I}_p^n\right)}{\mathbf{A}\left(\mathbf{R}_1^{\mathcal{L}^a}\right)}$$

$$= \mathbf{P}_{\mathcal{L}^a,\mathbf{I}^n}(i) - \frac{\sum_{p \in \mathbf{R}_1^{\mathcal{L}^a} \cap \mathbf{R}_0^{\mathcal{L}}} K_i\left(\mathbf{I}_p^n\right)}{\mathbf{A}\left(\mathbf{R}_1^{\mathcal{L}^a}\right)} \tag{A.4}$$

Using this lower bound in the Bhattacharyya measure, we obtain the following upper bound of $\mathbf{B}_{\mathbf{C},\mathbf{I}^n}(\mathcal{L})$ (i is omitted as argument of the distributions to simplify the equations)

$$\mathbf{B}_{\mathbf{C},\mathbf{I}^n}(\mathcal{L}) \leqslant -\sum_{i \in \mathcal{I}} \sqrt{\left(\mathbf{P}_{\mathcal{L}^a,\mathbf{I}^n} - \frac{\sum_{p \in \mathbf{R}_1^{\mathcal{L}^a} \cap \mathbf{R}_0^{\mathcal{L}}} K_i\left(\mathbf{I}_p^n\right)}{\mathbf{A}\left(\mathbf{R}_1^{\mathcal{L}^a}\right)}\right) \mathcal{M}_{\mathbf{C},\mathcal{I}}}$$

$$= -\sum_{i \in \mathcal{I}} \sqrt{\mathbf{P}_{\mathcal{L}^a,\mathbf{I}^n} \mathcal{M}_{\mathbf{C},\mathcal{I}}} \sqrt{1 - \frac{\sum_{p \in \mathbf{R}_1^{\mathcal{L}^a} \cap \mathbf{R}_0^{\mathcal{L}}} K_i\left(\mathbf{I}_p^n\right)}{\mathbf{P}_{\mathcal{L}^a,\mathbf{I}^n} \mathbf{A}\left(\mathbf{R}_1^{\mathcal{L}^a}\right)}}$$

$$= -\sum i \in \mathcal{I} \sqrt{\mathbf{P}_{\mathcal{L}^a,\mathbf{I}^n} \mathcal{M}_{\mathbf{C},\mathcal{I}}} \sqrt{1 - \frac{\sum_{p \in \mathbf{R}_1^{\mathcal{L}^a} \cap \mathbf{R}_0^{\mathcal{L}}} K_i\left(\mathbf{I}_p^n\right)}{\sum_{p \in \mathbf{R}_1^{\mathcal{L}^a}} K_i\left(\mathbf{I}_p^n\right)}} \tag{A.5}$$

Now notice the following inequality for any $0 \leqslant \mathbf{x} \leqslant 1$

$$\sqrt{1 - \mathbf{x}} \geqslant 1 - \mathbf{x} \tag{A.6}$$

Because $\mathbf{R}_1^{\mathcal{L}^a} \cap \mathbf{R}_0^{\mathcal{L}} \subset \mathbf{R}_1^{\mathcal{L}^a}$, we have

$$0 \leqslant \frac{\sum_{p \in \mathbf{R}_1^{\mathcal{L}^a} \cap \mathbf{R}_0^{\mathcal{L}}} K_i\left(\mathbf{I}_p^n\right)}{\sum_{p \in \mathbf{R}_1^{\mathcal{L}^a}} K_i\left(\mathbf{I}_p^n\right)} \leqslant 1 \tag{A.7}$$

Thus, applying inequality (A.6) gives

$$\sqrt{1 - \frac{\sum_{p \in \mathbf{R}_1^{\mathcal{L}^a} \cap \mathbf{R}_0^{\mathcal{L}}} K_i\left(\mathbf{I}_p^n\right)}{\sum_{p \in \mathbf{R}_1^{\mathcal{L}^a}} K_i\left(\mathbf{I}_p^n\right)}} \geqslant 1 - \frac{\sum_{p \in \mathbf{R}_1^{\mathcal{L}^a} \cap \mathbf{R}_0^{\mathcal{L}}} K_i\left(\mathbf{I}_p^n\right)}{\sum_{p \in \mathbf{R}_1^{\mathcal{L}^a}} K_i\left(\mathbf{I}_p^n\right)} \tag{A.8}$$

Finally, combining this inequality with (A.5) gives the following upper bound of $\mathbf{B}_{\mathbf{C},\mathbf{I}^n}(\mathcal{L})$

$$\mathbf{B}_{\mathbf{C},\mathbf{I}^n}(\mathcal{L}) \leqslant -\sum_{i \in \mathcal{I}} \sqrt{\mathbf{P}_{\mathcal{L}^a,\mathbf{I}^n} \mathcal{M}_{\mathbf{C},\mathcal{I}}} \left(1 - \frac{\sum_{p \in \mathbf{R}_1^{\mathcal{L}^a} \cap \mathbf{R}_0^{\mathcal{L}}} K_i\left(\mathbf{I}_p^n\right)}{\sum_{p \in \mathbf{R}_1^{\mathcal{L}^a}} K_i\left(\mathbf{I}_p^n\right)}\right)$$

$$= \mathbf{B}_{\mathbf{C},\mathbf{I}^n}(\mathcal{L}^a) + \sum_{i \in \mathcal{I}} \sqrt{\mathbf{P}_{\mathcal{L}^a,\mathbf{I}^n} \mathcal{M}_{\mathbf{C},\mathcal{I}}} \frac{\sum_{p \in \mathbf{R}_1^{\mathcal{L}^a} \cap \mathbf{R}_0^{\mathcal{L}}} K_i\left(\mathbf{I}_p^n\right)}{\mathbf{P}_{\mathcal{L}^a,\mathbf{I}^n} \mathbf{A}(\mathbf{R}_1^{\mathcal{L}^a})}$$

$$= \mathbf{B}_{\mathbf{C},\mathbf{I}^n}(\mathcal{L}^a) + \sum_{p \in \mathbf{R}_0^{\mathcal{L}}} \frac{\mathcal{L}_p^a}{\mathbf{A}\left(\mathbf{R}_1^{\mathcal{L}^a}\right)} \sum_{i \in \mathcal{I}} K_i\left(\mathbf{I}_p^n\right) \sqrt{\frac{\mathcal{M}_{\mathbf{C},\mathcal{I}}(i)}{\mathbf{P}_{\mathcal{L}^a,\mathbf{I}^n}(i)}} \tag{A.9}$$

Now because $\mathbf{A}\left(\mathbf{R}_1^{\mathcal{L}^a} \cap \mathbf{R}_0^{\mathcal{L}}\right) + \mathbf{A}(\mathbf{R}_1^{\mathcal{L}}) = \mathbf{A}\left(\mathbf{R}_1^{\mathcal{L}^a}\right)$, we have:

$$\mathbf{B_{C,I^n}}(\mathcal{L}^a) = \frac{\mathbf{B_{C,I^n}}(\mathcal{L}^a)\mathbf{A}\left(\mathbf{R}_1^{\mathcal{L}^a} \cap \mathbf{R}_0^{\mathcal{L}}\right)}{\mathbf{A}\left(\mathbf{R}_1^{\mathcal{L}^a}\right)} + \frac{\mathbf{B_{C,I^n}}(\mathcal{L}^a)\mathbf{A}\left(\mathbf{R}_1^{\mathcal{L}}\right)}{\mathbf{A}\left(\mathbf{R}_1^{\mathcal{L}^a}\right)} \tag{A.10}$$

Now notice the following equalities:

$$\mathbf{A}\left(\mathbf{R}_1^{\mathcal{L}^a} \cap \mathbf{R}_0^{\mathcal{L}}\right) = \sum_{p \in \mathbf{R}_0^{\mathcal{L}}} \mathcal{L}_p^a$$

$$\mathbf{A}(\mathbf{R}_1^{\mathcal{L}}) = \sum_{p \in \mathbf{R}_1^{\mathcal{L}}} 1 \tag{A.11}$$

Embedding these equalities in (A.10) gives:

$$\mathbf{B_{C,I^n}}(\mathcal{L}^a) = \sum_{p \in \mathbf{R}_0^{\mathcal{L}}} \frac{\mathcal{L}_p^a \mathbf{B_{C,I^n}}(\mathcal{L}^a)}{\mathbf{A}\left(\mathbf{R}_1^{\mathcal{L}^a}\right)} + \sum_{p \in \mathbf{R}_1^{\mathcal{L}}} \frac{\mathbf{B_{C,I^n}}(\mathcal{L}^a)}{\mathbf{A}\left(\mathbf{R}_1^{\mathcal{L}^a}\right)} \tag{A.12}$$

Now because the last term in (A.12) is positive, we have the following inequality $\forall \alpha \in [0,1]$:

$$\mathbf{B_{C,I^n}}(\mathcal{L}^a) \leqslant \sum_{p \in \mathbf{R}_0^{\mathcal{L}}} \frac{\mathcal{L}_p^a \mathbf{B_{C,I^n}}(\mathcal{L}^a)}{\mathbf{A}\left(\mathbf{R}_1^{\mathcal{L}^a}\right)} + (1-\alpha) \sum_{p \in \mathbf{R}_1^{\mathcal{L}}} \frac{\mathbf{B_{C,I^n}}(\mathcal{L}^a)}{\mathbf{A}\left(\mathbf{R}_1^{\mathcal{L}^a}\right)} \tag{A.13}$$

Combining (A.9) and (A.13) proves the inequality in Eq. (14). Proof of the inequality in Eq. (16) follows the same steps.

Appendix B. Proof of Proposition 2

To prove Proposition 2, it suffices to verify conditions (18) and (19) for $\mathcal{A_{C,I^n}}$ and $\mathcal{F_{C,I^n}}$. Condition (18) follows directly from Proposition 1. For Condition (19), it suffices to see that

When $\mathcal{L}^a = \mathcal{L}, \mathcal{L}_p^a = 0 \quad \forall p \in \mathbf{R}_0^{\mathcal{L}}$, i.e., $\begin{cases} \mathbf{c}_{p,\mathbf{I}^n}(\mathbf{0}) = 0 \\ \mathbf{c}_{p,\mathbf{D}}(\mathbf{0}) = 0 \end{cases} \quad \forall p \in \mathbf{R}_0^{\mathcal{L}}$

In this case, we have

$$\sum_{p \in \mathbf{R}_0^{\mathcal{L}}} \mathbf{c}_{p,\mathbf{I}^n}(\mathbf{0}) = 0 \quad \text{and} \quad \sum_{p \in \mathbf{R}_0^{\mathcal{L}}} \mathbf{c}_{p,\mathbf{D}}(\mathbf{0}) = 0 \tag{B.1}$$

Therefore,

for $\alpha = 0$, $\quad \mathcal{A_{C,I^n}}(\mathcal{L}, \mathcal{L}, 0)$

$$= \mathcal{J_{C,I^n}}(\mathcal{L}, \mathcal{L}, 0) + \gamma \mathcal{J_{C,D}}(\mathcal{L}, \mathcal{L}, 0) + \lambda \mathbf{S}(\mathcal{L})$$

$$= \sum_{p \in \mathbf{R}_1^{\mathcal{L}}} \mathbf{c}_{p,\mathbf{I}^n}(\mathbf{1}) + \gamma \sum_{p \in \mathbf{R}_1^{\mathcal{L}}} \mathbf{c}_{p,\mathbf{D}}(\mathbf{1}) + \lambda \mathbf{S}(\mathcal{L})$$

$$= \sum_{p \in \mathbf{R}_1^{\mathcal{L}}} \frac{\mathbf{B_{C,I^n}}(\mathcal{L})}{\mathbf{A}(\mathbf{R}_1^{\mathcal{L}})} + \gamma \sum_{p \in \mathbf{R}_1^{\mathcal{L}}} \frac{\mathbf{B_{C,D}}(\mathcal{L})}{\mathbf{A}(\mathbf{R}_1^{\mathcal{L}})} + \lambda \mathbf{S}(\mathcal{L})$$

$$= \mathbf{B_{C,I^n}}(\mathcal{L}) + \gamma \mathbf{B_{C,D}}(\mathcal{L}) + \lambda \mathbf{S}(\mathcal{L}) = \mathcal{F_{C,I^n}}(\mathcal{L}) \tag{B.2}$$

which verifies condition (19).

References

Adam, A., Kimmel, R., Rivlin, E., 2009. On scene segmentation and histograms-based curve evolution. IEEE Transactions on Pattern Analysis and Machine Intelligence 31, 1708–1714.

Aherne, F., Thacker, N., Rockett, P., 1997. The Bhattacharyya metric as an absolute similarity measure for frequency coded data. Kybernetika 32, 1–7.

Andreopoulos, A., Tsotsos, J.K., 2008. Efficient and generalizable statistical models of shape and appearance for analysis of cardiac MRI. Medical Image Analysis 12, 335–357.

Aubert, G., Barlaud, M., Faugeras, O., Jehan-Besson, S., 2003. Image segmentation using active contours: calculus of variations or shape gradients? SIAM Journal of Applied Mathematics 63, 2128–2154.

Ben Ayed, I., Chen, H.M., Punithakumar, K., Ross, I., Li, S., 2010. Graph cut segmentation with a global constraint: recovering region distribution via a bound of the Bhattacharyya measure. In: IEEE International Conference on Computer Vision and Pattern Recognition (CVPR 2010), San Francisco, CA, USA, pp. 1–7.

Ben Ayed, I., Li, S., Ross, I., 2009a. Embedding overlap priors in variational left ventricle tracking. IEEE Transactions on Medical Imaging 28, 1902–1913.

Ben Ayed, I., Li, S., Ross, I., 2009b. A statistical overlap prior for variational image segmentation. International Journal of Computer Vision 85, 115–132.

Ben Ayed, I., Li, S., Ross, I., Islam, A., 2009c. Myocardium tracking via matching distributions. International Journal of Computer Assisted Radiology and Surgery 4, 37–44.

Ben Ayed, I., Punithakumar, K., Li, S., Islam, A., Chong, J., 2009. Left ventricle segmentation via graph cut distribution matching. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI 2009), vol. 1, London, UK, pp. 901–909.

Bishop, C.M., 2007. Pattern Recognition and Machine Learning (Information Science and Statistics). Springer.

Blake, A., Rother, C., Brown, M., Pérez, P., Torr, P.H.S., 2004. Interactive image segmentation using an adaptive gmmrf model. In: European Conference on Computer Vision (ECCV 2004), vol. 1, Prague, Czech Republic, pp. 428–441.

Boykov, Y., Funka Lea, G., 2006. Graph cuts and efficient n-d image segmentation. International Journal of Computer Vision 70, 109–131.

Boykov, Y., Kolmogorov, V., 2003. Computing geodesics and minimal surfaces via graph cuts. In: IEEE International Conference on Computer Vision (ICCV 2003), vol. 1, Nice, France, pp. 26–33.

Boykov, Y., Kolmogorov, V., 2004. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. IEEE Transactions on Pattern Analysis and Machine Intelligence 26, 1124–1137.

Cerqueira, M.D., Weissman, N.J., D, V., Jacobs, A.K., Kaul, S., Laskey, W.K., Pennell, D.J., Rumberger, J.A., Ryan, T., Verani, M.S., 2002. Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart: a statement for healthcare professionals from cardiac imaging committee of the council on clinical cardiology of the american heart association. Circulation 105, 539–542.

Comaniciu, D., Ramesh, V., Meer, P., 2003. Kernel-based object tracking. IEEE Transactions on Pattern Analysis and Machine Intelligence 25, 564–575.

El-Berbari, R., Bloch, I., Redheuil, A., Angelini, E.D., Mousseaux, É., Frouin, F., Herment, A., 2007. Automated segmentation of the left ventricle including papillary muscles in cardiac magnetic resonance images. In: Functional Imaging and Modeling of the Heart (FIMH 2007), Salt Lake City, UT, USA, pp. 453–462.

Flusser, J., Zitova, B., Suk, T., 2009. Moments and Moment Invariants in Pattern Recognition. Wiley.

Fradkin, M., Ciofolo, C., Mory, B., Hautvast, G., Breeuwer, M., 2008. Comprehensive segmentation of cine cardiac mr images. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI 2008), vol. 1, New York, NY, USA, pp. 178–185.

Freedman, D., Zhang, T., 2004. Active contours for tracking distributions. IEEE Transactions on Image Processing 13, 518–526.

Fritscher, K.D., Pilgram, R., Schubert, R., 2005. Automatic cardiac 4d segmentation using level sets. In: Functional Imaging and Modeling of the Heart (FIMH 2005), Barcelona, Spain, pp. 113–122.

Georgiou, T., Michailovich, O., Rathi, Y., Malcolm, J., Tannenbaum, A., 2007. Distribution metrics and image segmentation. Linear Algebra and its Applications 425, 663–672.

Hautvast, G., Lobregt, S., Breeuwer, M., Gerritsen, F., 2006. Automatic contour propagation in cine cardiac magnetic resonance images. IEEE Transactions on Medical Imaging 25, 1472–1482.

Jolly, M.P., 2008. Automatic recovery of the left ventricular blood pool in cardiac cine mr images. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI 2008), vol. 1, New York, NY, USA, pp. 110–118.

Juan, O., Boykov, Y., 2007. Capacity scaling for graph cuts in vision. In: IEEE International Conference on Computer Vision (ICCV 2007).

Kaus, M.R., von Berg, J., Weese, J., Niessen, W., Pekar, V., 2004. Automated segmentation of the left ventricle in cardiac mri. Medical Image Analysis 8, 245–254.

Kohli, P., Rihan, J., Bray, M., Torr, P.H.S., 2008. Simultaneous segmentation and pose estimation of humans using dynamic graph cuts. International Journal of Computer Vision 79, 285–298.

Lee, D.D., Seung, H., 2000. Algorithms for nonnegative matrix factorization. In: Neural Information Processing Systems (NIPS 2000), vol. 13, Denver, CO, USA, pp. 556–562.

Lehmann, E.L., 1986. Testing Statistical Hypotheses. Wiley, New York.

Liu, H., Chen, Y., Ho, H.P., Shi, P., 2005. Geodesic active contours with adaptive neighboring influence. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI 2005), vol. 2, Palm Springs, CA, USA, pp. 741–748.

Lorenzo-Valdés, M., Sanchez-Ortiz, G.I., Elkington, A.G., Mohiaddin, R.H., Rueckert, D., 2004. Segmentation of 4d cardiac mr images using a probabilistic atlas and the em algorithm. Medical Image Analysis 8, 255–265.

Lynch, M., Ghita, O., Whelan, P., 2008. Segmentation of the left ventricle of the heart in 3-D+T MRI data using an optimized nonrigid temporal model. IEEE Transactions on Medical Imaging 27, 195–203.

Michailovich, O.V., Rathi, Y., Tannenbaum, A., 2007. Image segmentation using active contours driven by the Bhattacharyya gradient flow. IEEE Transactions on Image Processing 16, 2787–2801.

Mitchell, S.C., Bosch, J.G., Lelieveldt, B.P.F., van der Geest, R.J., Reiber, J.H.C., Sonka, M., 2002. 3-D active appearance models: segmentation of cardiac mr and ultrasound images. IEEE Transactions on Medical Imaging 21, 1167–1178.

Mitchell, S.C., Lelieveldt, B.P.F., van der Geest, R.J., Bosch, J.G., Reiber, J.H.C., Sonka, M., 2001. Multistage hybrid active appearance model matching: segmentation of left and right ventricles in cardiac mr images. IEEE Transactions on Medical Imaging 20, 415–423.

Ni, K., Bresson, X., Chan, T.F., Esedoglu, S., 2009. Local histogram based segmentation using the wasserstein distance. International Journal of Computer Vision 84, 97–111.

Pluempitiwiriyawej, C., Moura, J.M.F., Wu, Y.J.L., Ho, C., 2005. Stacs: new active contour scheme for cardiac mr image segmentation. IEEE Transactions on Medical Imaging 24, 593–603.

Punithakumar, K., Ben Ayed, I., Islam, A., Ross, I., Li, S., 2010. Tracking endocardial motion via multiple model filtering. IEEE Transactions on Biomedical Engineering 57, 2001–2010.

Rother, C., Kolmogorov, V., Blake, A., 2004. Grabcut: interactive foreground extraction using iterated graph cuts. ACM Transactions on Graphics 23, 309–314.

Rother, C., Minka, T.P., Blake, A., Kolmogorov, V., 2006. Cosegmentation of image pairs by histogram matching - incorporating a global constraint into mrfs. In: IEEE International Conference on Computer Vision and Pattern Recognition (CVPR 2006), vol. 1, New York, NY, USA, pp. 993–1000.

Rousson, M., Cremers, D., 2005. Efficient kernel density estimation of shape and intensity priors for level set segmentation. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI 2005), vol. 2, Palm Springs, CA, USA, pp. 757–764.

Spottiswoode, B.S., Zhong, X., Lorenz, C.H., Mayosi, B.M., Meintjes, E.M., Epstein, F.H., 2009. Motion-guided segmentation for cine dense mri. Medical Image Analysis 13, 105–115.

Sun, W., Cetin, M., Chan, R., Reddy, V., Holmvang, G., Chandar, V., Willsky, A.S., 2005. Segmenting and tracking the left ventricle by learning the dynamics in cardiac images. In: Information Processing in Medical Imaging (IPMI 2005), Glenwood Springs, CO, USA, pp. 553–565.

Van Assen, H.C., Danilouchkine, M.G., Dirksen, M.S., Reiber, J.H.C., Lelieveldt, B.P.F., 2008. A 3-d active shape model driven by fuzzy inference: application to cardiac ct and mr. IEEE Transactions on Information Technology in Biomedicine 12, 595–605.

Zambal, S., Hladuvka, J., Bühler, K., 2006. Improving segmentation of the left ventricle using a two-component statistical model. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI 2006), vol. 1, Copenhagen, Denmark, pp. 151–158.

Zhang, T., Freedman, D., 2005. Improving performance of distribution tracking through background mismatch. IEEE Transactions on Pattern Analysis and Machine Intelligence 27, 282–287.

Zhu, Y., Papademetris, X., Sinusas, A.J., Duncan, J.S., 2010. Segmentation of the left ventricle from cardiac mr images using a subject-specific dynamical model. IEEE Transactions on Medical Imaging 29, 669–687.

Zhu-Jacquot, J., Zabih, R., 2008. Segmentation of the left ventricle in cardiac mr images using graph cuts with parametric shape priors. In: IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2008), Caesars Palace, Las Vegas, Nevada, USA, pp. 521–524.

Zhuang, X., Rhode, K.S., Arridge, S.R., Razavi, R., Hill, D.L.G., Hawkes, D.J., Ourselin, S., 2008. An atlas-based segmentation propagation framework using locally affine registration - application to automatic whole heart segmentation. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI 2008), vol. 2, New York, NY, USA, pp. 425–433.