J Digit Imaging (2013) 26:898–908
DOI 10.1007/s10278-013-9573-z

Cardiac MRI Segmentation Using Mutual Context Information from Left and Right Ventricle

Dwarikanath Mahapatra

Published online: 25 January 2013

Abstract In this paper, we propose a graphcut method to segment the cardiac right ventricle (RV) and left ventricle (LV) by using context information from each other. Contextual information is very helpful in medical image segmentation because the relative arrangement of different organs is the same. In addition to the conventional log-likelihood penalty, we also include a "context penalty" that captures the geometric relationship between the RV and LV. Contextual information for the RV is obtained by learning its geometrical relationship with respect to the LV. Similarly, RV provides geometrical context information for LV segmentation. The smoothness cost is formulated as a function of the learned context which helps in accurate labeling of pixels. Experimental results on real patient datasets from the STACOM database show the efficacy of our method in accurately segmenting the LV and RV. We also conduct experiments on simulated datasets to investigate our method's robustness to noise and inaccurate segmentations.

Keywords Mutual context information · LV · RV · Segmentation · Cardiac · MRI · Graph cuts

Introduction

Cardiovascular diseases are one of the leading causes of death in the Western world [1]. Diagnosis and treatment relies on numerous imaging modalities like echography, computed tomography (CT), coronary angiography, and magnetic resonance imaging (MRI). MRI has emerged as the preferred diagnostic modality because of its non-invasive nature. It also gives reliable information on morphology, muscle perfusion, tissue viability, and blood flow. These parameters are obtained by segmenting the left ventricle (LV) and right ventricle (RV) from cardiac MR images. Manual segmentation is tedious and prone to intra- and inter-observer variability. This has necessitated the development of automated/semiautomated segmentation algorithms. An exhaustive review of medical image segmentation algorithms can be found in [2], while an excellent review of cardiac LV segmentation algorithms is given in [3]. While there are many methods for LV segmentation, the RV has not received so much attention [3] because: (1) of its complex crescent shape, (2) lower pressure to eject blood, (3) thinner structure than LV, and (4) less critical function than LV. However, Shors et al. [4] show that MRI can provide an accurate quantification of RV mass.

D. Mahapatra (✉)
Department of Computer Science, Swiss Federal Institute of Technology (ETH) Zurich, Room CAB F 61.1 Universitätstrasse, 68092 Zurich, Switzerland
e-mail: dmahapatra@gmail.com

Knowledge-Based Cardiac Segmentation

Most cardiac MR images show poor contrast between LV blood pool and myocardium wall, thus giving minimal edge information. This, in addition to similar intensity distributions in different regions, makes segmentation of the LV a very challenging task when using only low level information (e.g., intensity, gradient, etc.). RV segmentation poses challenges because of the reduced thickness in MRI and shape variations. In such a scenario, inclusion of prior shape information assumes immense significance in LV and RV segmentation. Although there is a vast body of literature for cardiac LV segmentation from short-axis images, we only review those works that incorporate different kinds of prior knowledge into the segmentation task.

Jolly propose a LV blood pool localization approach in [5] which acts as an initialization for LV segmentation. Paragios et al. [6] used a signed distance map to incorporate prior shape knowledge in a level set framework for LV segmentation. A probability density function of the shape

or appearance from training data was used in level sets [7] and graph cuts [8], while orientation histograms were used as shape priors for graphcut segmentation of the LV [9, 10]. Cousty et al. [11] used mixture models and spatiotemporal watershed transforms to segment the LV while [12] employed clustering approaches. Ayed et al. employed level sets to match overlap priors which is the relation between overlapping intensity distributions of different areas like myocardium and blood pool. Mutually beneficial registration and segmentation information was combined for LV segmentation in [13–16].

Kaus et al. in [17] combine a statistical model with coupled mesh surfaces for segmentation. However, their assumption that the heart is located in the center of the image is not always valid. Shape knowledge was combined with dynamic information to account for cardiac shape variability in [18, 19]. Zhu et al. in [18] use a subject specific dynamic model that handles intersubject variability and temporal dynamics (intrasubject) variability simultaneously. A recursive Bayesian framework is then used for segmenting each frame. In [19], the cardiac dynamics is learnt by using a second-order dynamic model. Davies et al. in [20] propose a method to automatically extract a set of optimal landmarks using the minimum description length. But it is not clear whether the landmarks thus extracted are optimal in the sense of anatomical correspondence. Perperidis et al. proposed a 4-D model by including temporal information [21]. Besbes et al. [22] used a control point representation of the LV prior and other images were deformed to match the shape prior.

Active appearance models (AAM) were used in [23, 24] to segment the LV. AAM and ASM (active shape models) were combined in [25], with extensions to the time domain in [26]. Some of the works on LV segmentation also show results for RV segmentation using deformable models [27, 28] and atlas-based methods [29]. However, to our knowledge, none of the works use any kind of contextual information from the RV (or LV) to segment the LV (or RV). Our work uses a novel approach to model the contextual relationship between the RV and LV and achieve better segmentation accuracy.

Context-Based Segmentation

Use of contextual information is extremely popular in general computer vision tasks like segmentation and scene understanding. Markov random fields (MRFs) have been widely used in vision [30] along with conditional MRFs (CRFs) [31]. These models have similar disadvantages like: (1) limited choice of functions, (2) they rely on a fixed neighborhood giving very limited topological information, and (3) obtain the optimal solution for a limited family of functions. Many important works have been proposed to include context information for segmentation. Belongie et al. [32] proposed a shape context method using distribution of relative distances for matching shapes. Hoiem et al. in [33] learn the relationship between different target objects for locating pedestrians, cars, and buildings from natural images. Other works have modeled the direct relationship between different objects [31], regions [34], or scene categories [35]. In [36], Tu propose "auto-context" which integrates image and contextual information from a set of trained classifiers to segment brain structures from MRI. Auto context was used by Li et al. in [37] to segment the human prostate gland from CT images.

Graph cut-based multiregion segmentations were introduced in [38] where multiple regions were segmented simultaneously. Object interaction prior modeling the geometric relationship between different disk-like structures were used in [39] for intervertebral disk segmentation using graph cuts. Based on the segmentation of one disk, other interesting structures are segmented. Song et al. [40] proposed a surface region context model for segmenting pulmonary tumors. Context comes from a variety of sources. Since the human anatomy is standard, and image acquisition procedures are the same, presence of one organ gives a strong cue about the presence of other organs in the medical images. For example in abdominal MRI, the spine can be easily identified, while segmenting other organs like the kidney, liver, or prostate is challenging. It can be argued that first by segmenting the spine we can include contextual information to aid in the segmentation of other organs. A similar scenario is encountered in segmenting the RV and LV. Both of them provide mutually complementary contextual cues to aid in segmentation.

Our Contribution

We propose a method to segment the RV and LV by learning their mutual context information. None of the previous works model the contextual relationship between RV and LV for their mutual segmentation. We make use of geometric relationship between the shapes in the form of relative orientations. The context information is learned from a set of training images in which the RV and LV have been manually segmented. The learned information is encoded on a graphical model of the image as weights between pixel nodes and terminal nodes, as well as the weights between pixel nodes. Graph cuts are used to find the final labels in an iterative fashion. Geometric relationships in the form of orientation angles and distance were used in [39]. However, they try to match distributions of different areas, and hence do not encode costs at levels of pixels or their neighborhoods. Therefore, graph cuts cannot be directly used for their formulation. The authors introduce submodular graphs to overcome this limitation and optimize over a upper bound which enables use of graph cuts for optimization.

This paper makes the following contributions: (1) we encode context information for each pixel as well as between pixels. This allows us to directly use graph cuts without constructing submodular graphs as in [39]; (2) mutually beneficial context information from the RV and LV is used to segment the individual organs. We describe our method in "Materials and Methods" section, present experimental results in "Experiments and Results" section, and conclude with "Conclusion" section.

Materials and Methods

Overview of Method

Our method comprises of the following steps: (1) select seed points on the RV and LV, and use graph cuts to get initial segmentations; (2) fix the LV segmentation and refine the RV segmentation using contextual information from the LV; (3) fix the new RV segmentation and refine LV segmentation using contextual information from the RV; (4) repeat steps 2 and 3 till there is no further change in both segmentations.

Initial Segmentation of the RV and LV

Every segmentation task has to be defined within a segmentation framework of a particular cost function. Every cost function has two components: (1) the penalty cost—which quantifies the deviation from the ideal labels and (2) smoothness cost that regularizes the solution so that neighboring pixels have similar segmentation labels. We use the second-order MRF formulation to define our segmentation costs. The cost function is written as

$$E(L) = \sum_{s \varepsilon P} D_I(L_s) + \lambda \sum_{(s,t) \varepsilon N} V_I(L_s, L_t), \tag{1}$$

where P denotes the set of pixels, L_s is the label of pixel $s \in P$ and denotes its segmentation class, and N is the set of neighboring pixel pairs. $D_I(L_s)$ is the data penalty function that measures how well label L_s fits pixel s; V_I is the smoothness cost that measures the cost of assigning labels L_s and L_t to neighboring pixels s and t. Subscript I denotes intensity. λ is a weight that determines the relative contribution of the two terms. Eq. 1 is optimized using graph cuts [41].

To calculate the penalty cost of a pixel for a particular class, we need to have a reference function for each class. For the segmentation task, the reference function is the intensity distribution of pixel intensities known to belong to a particular class. Pixels from different classes (known as seed points) are chosen from an image with known labels, and their intensity distributions are modeled as a Gaussian distribution. The penalty function $D_I(L_s)$ is defined as the negative log-likelihood of intensity:

$$D_I(L_s) = -\log \Pr(I_s / L_s) \tag{2}$$

where I_s is the intensity at pixel s, Pr is the likelihood (obtained by inserting the intensity value into the reference distribution for class L), and L_s is the label. V_I assigns a low penalty at edge points to ensure neighboring pixels have similar labels, and favors a piecewise constant segmentation result. It is defined as

$$V_I(L_s, L_t) = \begin{cases} e^{-\frac{(I_s - I_t)^2}{2\sigma^2}} \times \frac{1}{\|s-t\|}, & L_s \neq L_t \\ 0 & L_s = L_t \end{cases} \tag{3}$$

σ determines the intensity difference up to which a region is considered as piecewise smooth. It is equal to the average intensity difference in a neighborhood of s. $\|s-t\|$ is the Euclidean distance between pixels s and t. I_s and I_t are the intensities of neighboring pixels s and t. By optimizing Eq. 1, we get a set of initial LV and RV segmentations which are not optimal because intensity information alone is unable to overcome the challenges of low resolution noisy images and indistinct edges. We refine the LV and RV segmentations by incorporating contextual information which is described in the subsequent sections.

Modeling Contextual Relationship between RV and LV

To model the contextual relationship between the RV and LV, we use 40 training images in which the RV and LV have been manually segmented. The STACOM dataset that we use provides only the manual segmentation of the LV. The RV is segmented by two radiologists each having more than 5 years of clinical experience (details are given in the "Description of Datasets" section). We shall denote these shapes as Rv and Lv. Let the set of points on the outer edge of Rv be denoted as $\{Rv_1, \cdots, Rv_i\}$, and the set of points on Lv is denoted as $\{Lv_1; \cdots; Lv_j\}$. We choose points only on the outer edge as they are sufficient to give a rich shape descriptor without large computational overhead.

Figure 1b shows an illustration of our approach to model the contextual relationship between the LV and RV. We calculate the histogram of orientation angles (i.e., the distribution of orientation angles) between each Lv_j and all Rv_i (θ_{LV} in Fig. 1b) and denote it as $h^o{}_{Lv,j}$. The superscript o denotes orientation angles. The 32 histogram bins are uniform in log-polar space making the descriptor more sensitive to positions of nearby sample points than points further away. Figure 1c shows $h^o{}_{Lv,j}$ for the manual segmentations (shown in red) in Fig. 1a. The angles are distributed over different range of values and have more than one mode in most cases. The x-axis gives the angle value in degrees and is in the range $[-180°, 180°]$. If we consider all the Lv_j over

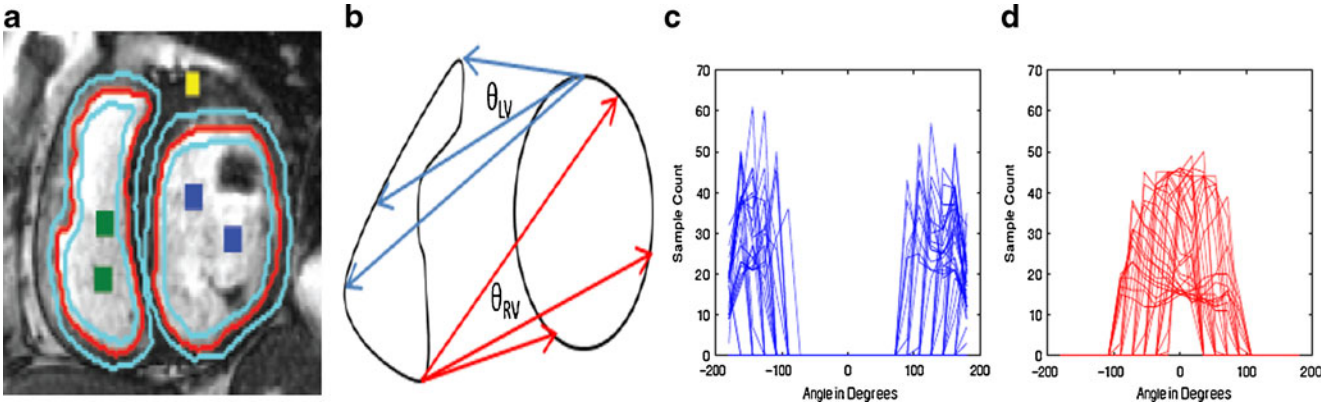

Fig. 1 Illustrations for context segmentation. **a** Example image with the manual segmentations in *red* and the corresponding narrow bands in *cyan*. **b** Illustration of contextual relationship between LV and RV. **c** Angle distribution of LV with respect to RV. **d** Angle distribution for RV with respect to LV

all training images, then there are a large number of $h^{o}_{\mathrm{Lv,j}}$. The representative distributions are obtained by principal component analysis (PCA). PCA gives the principal modes of the distribution, i.e., those modes along which most of the data is distributed. The principal components of the first six histograms represents more than 94 % energy (as determined from the Eigenvalues) and are chosen as the reference histograms for determining the context penalty. These histograms are denoted as $H^{o}_{\mathrm{Lv},n}$, $n=\{1, \cdots, 6\}$. Thus, we obtain the contextual relationship of the LV with respect to the RV which is the relationship of orientation angles between each point on the LV with respect to all points on the RV.

Similarly, the contextual relationship of the RV with respect to the LV means the relationship of orientation angles between each point on the RV with respect to all points on the LV. This relationship is obtained by calculating the histogram of orientation angles between each Rvi and all Lvj and denoted as $h^{o}_{\mathrm{Rv,i}}$. Note the difference in notation (i.e., Rv and i). The six principal histograms are denoted by $H^{o}_{\mathrm{Rv},n}$, $n=\{1, \cdots, 6\}$. Since contextual information is incorporated to refine the segmentations obtained in the initial step we choose to calculate the cost function values over a narrow band around the segmented RV or LV. The narrow band is obtained by expanding the contour outline (obtained in the initial segmentation steps) in both directions normal to its edges. The expanded distance is 20 pixels in each direction and the width of the band is 40 pixels. Figure 1a shows the narrow band (in cyan) around which the pixel labels are updated.

For a pixel in the band around the LV, let its orientation histogram with respect to the Rv be denoted as h^{o}. Distance of h^{o} from each $H^{o}_{\mathrm{Lv},n}$ is given by the χ^2 metric as

$$d_L V(o,n) = \frac{1}{2}\sum_{(k=1)}^{K} \frac{\left[h^{\circ}(k) - H^{\circ}_{(Lv,n)(k)}\right]^2}{h^{\circ}(k) + H^{\circ}_{(Lv,n)(k)}(k)}, \tag{4}$$

Where k denotes the bin index of the histogram and $K=32$ is the total number of bins. Note that the distance is calculated for all $n=6$ reference distributions of the LV. The $\chi\partial$ metric gives values between 0 (identical distributions) and 1 (no similarity in distribution), although other metrics (like Bhattacharya distance) can also be used. The "context cost" (D_{Cont}) of assigning label 1 (LV) to the pixel is the minimum of $d_{\mathrm{Lv}}(o,n)$

$$D_{\mathrm{Cont}}(L_s = 1) = \arg_{\mathrm{n}} \min \mathrm{d_{LV}(o,n)} \tag{5}$$

The principal component among $H^{o}_{\mathrm{Lv},n}$ giving the minimum cost represents the most similar direction of h^{o} in the training dataset and hence also the similarity of h^{o} with Lv. Therefore, it makes sense to take the minimum of all histogram distances. The context cost of assigning label 0 (background) to the pixel is defined as

$$D_{\mathrm{Cont}}(L_s = 0) = 1 - D_{\mathrm{Cont}}(L_s = 1) \tag{6}$$

Note that for each round of segmentations there are only two labels (RV or LV, and background). Thus the above formulation of $D_{\mathrm{cont}}(\mathrm{L_S}=0)$ is justified. The learned contextual relationship is incorporated into the smoothness cost by the following approach.

The pairwise cost is the weight between the two corresponding pixels (nodes) on the graph. If the two pixels have the same labels, then their mutual weight should be high such that the link is not severed (because graph cut optimization severs links in such a way that the combined sum of weights of severed links is minimal). On the other hand, if the two pixels have different labels then their edge weight should be low. For a neighboring pixel pair s and t around the LV, we determine their distribution of orientation angles with respect to the current RV segmentation, which we denote as $h^{o}{}_{s}$ and $h^{o}{}_{t}$. The difference in histograms is

calculated using the χ^2metric and denoted as $|h^o s,\ h^o t|$. The smoothness cost between s and t due to context is defined as

$$V_{\text{Cont}}(L_s, L_t) = 1 - |h_s, h_t| \tag{7}$$

Similar pixels s and t will have low value of $|h^o s,\ h^o t|$, and the corresponding V_{Cont} is high. Dissimilar neighboring pixels s and t will have high value of $|h^o s,\ h^o t|$, and the corresponding V_{Cont} is low. Thus, the above formulation of V_{Cont} serves our desired purpose. To obtain the updated segmentation, we incorporate both intensity penalty and contextual penalty in the cost function. Thus, the final energy function is

$$E(L) = \sum\nolimits_{s\varepsilon P} [D_I(L_s) + D_{Cont}(L_s)] + \lambda \sum\nolimits_{(s,t)\varepsilon N} [V_I(L_s, L_t) + V_{Cont}(L_s, L_t)] \tag{8}$$

where λ=0.05. Equation 8 is optimized using graph cuts to get an updated LV segmentation. The obtained LV segmentation is used to update the RV segmentation by steps similar to above. For a pixel in the band around the RV, its orientation histogram with respect to the Lv is calculated (h^o) and its distance from each $H^o{}_{\text{Rv},n}$ is given by

$$d_{RV}(o, n) = \frac{1}{2} \sum\nolimits_{k=1}^{K} \frac{\left[h^\circ(k) - H^\circ_{(\text{Rv,n})(\text{k})}\right]^2}{h^\circ(k) + H^\circ_{\text{Rv,n}}(k)}, \tag{9}$$

The D_{Cont} of assigning label 1 (RV) to the pixel is the minimum of $d_{\text{Rv}}(o; n)$,

$$D_{\text{Cont}}(L_s = 1) = \arg_{\text{n}}\ \min \text{d}_{\text{RV}}(\text{o}, \text{n}) \tag{10}$$

$D_{\text{Cont}}(L_s$=0) and $V_{\text{Cont}}(s,\ t)$ are calculated according to Eqs. 6 and 7, and the final energy function is obtained as in Eq. 8 which is optimized to get the updated RV segmentation. The sequence of steps is repeated till convergence.

Convergence Criteria The iterative segmentation is stopped if the change in segmentations between consecutive iterations is small. We calculate the dice metric (DM) between two consecutive segmentations. If the DM is greater than 95, then the segmentation is immediately stopped. The optimal segmentations are obtained in three to four iterations.

Optimization Using Graph Cuts

Pixels are represented as nodes Vp in a graph G which also consists of a set of directed edges E that connect two nodes. The edge weight between two neighboring nodes is the smoothness term while the data penalty term is the edge weight for links between nodes and label nodes (terminal nodes). The optimum labeling set is obtained by severing the edge links in such a manner that the cost of the cut is minimum. The number of nodes is equal to the number of pixels N and the number of labels is equal to L. There are two labels for each iteration, i.e., 0 for background and 1 for RV or LV. Details of graph construction and optimization can be found in [41].

Evaluation Criteria

The accuracy of the segmentation results against the ground truth was evaluated using the following metrics: sensitivity (p), specificity (q), positive predictive value (PPV), and negative predictive value (NPV). These were calculated by using the following equations:

$$p = \frac{T1}{N1},\quad q == \frac{T0}{N0},\ \text{PPV} = \frac{\text{T1}}{\text{T1} + \text{F1}},\ \text{NPV} = \frac{\text{T0}}{\text{To} + \text{F0}} \tag{11}$$

where $T1$ and $T0$ are the number of detected pixels characterized correctly as myocardium and nonmyocardium, while F1 and F0 are the number of misclassified pixels detected as myocardium and non-myocardium, respectively. The total number of myocardial and nonmyocardial pixels is $N1$ and $N0$, respectively. Other commonly used evaluation metrics include similarity indices in terms of the dice metric:

$$\text{DM}(\text{D1}, \text{T1}) = \frac{2|\text{D1} \cap \text{T1}|}{|\text{D1}| + |\text{T1}|} \tag{12}$$

and the Jaccard index:

$$JI(D1, T1) = \frac{|D1 \cap T1|}{|D1\ U\ T1|} \tag{13}$$

where $D1$ and $T1$ are automatic and ground truth segmentations, and $|X|$ denotes the number of elements in the set X. In both cases, values closer to 1 represent better performance.

Hausdorff Distance (HD) The DM gives a measure of how much the actual manual segmentation was recovered by the automatic segmentation. But the boundaries of the segmented regions may be far apart. The HD aims to measure the distance between the contours corresponding to different segmentations. We follow the definition of HD as given in [42]. If two curves are represented as sets of points A={a_1, a_2,…} and B={b_1, b_2,…}, where each a_i and b_i is an ordered pair of the x- and y-coordinates of a point on the curve, the distance to the closest point (DCP) of a_i to the curve B is calculated. The HD is defined as the maximum of the DCPs between the two curves [43]

$$\text{HD} = \max\left(\text{DCP}\{(\text{a}_\text{i}, \text{B})\}, \text{DCP}\left\{\left(\text{b}_\text{j}, \text{A}\right)\right\}\right) \tag{14}$$

Experiments and Results

Description of Datasets

We have tested our algorithm on 30 training datasets from the STACOM 2011 4D LV Segmentation Challenge run by the Cardiac Atlas Project [44]. All the data came from the Defibrillators to Reduce Risk by Magnetic Resonance Imaging Evaluation [45] cohort which consists of patients with coronary artery disease and prior myocardium infarction. The data were acquired using steady-state free precession MR imaging protocols with thickness of ≤10 mm, gap of ≤2 mm, TR of 30–50 ms, TE of 1.6 ms, flip angle of 60°, FOV of 360 mm, and 256×256 of image matrix. MRI parameters varied between cases, giving a heterogeneous mix of scanner types and imaging parameters consistent with typical clinical cases. The images consist of two to six LA slices and on average 12 SA slices and 26 time frames. Possible breath hold-related misalignments between different LA and SA slices were corrected using the method of [46]. The ground truth images of the LV were defined by an expert using an interactive guide point modeling segmentation algorithm [47]. Although there are 100 training datasets, none of them have manual segmentations of the RV. Two radiologists from the National University Hospital, Singapore, each having more than 5 years of clinical experience in studying patients with cardiac abnormalities, manually annotated the RV using he interactive guide point modeling segmentation algorithm [47]. The average dice metric between the two sets of contours is greater than 0.97 indicating good agreement between the individual segmentations. Therefore, we take the reference RV segmentation as the average of the two segmentations. The RV was segmented in 30 (out of 100) datasets.

The dataset was divided into training (15 patients) and test (15 patients) data. The training and test data was permuted to obtain 30 different groups such that each patient data was part of the training and test data. The reported results are the average of these 30 dataset groups. Automatic segmentations were obtained using three methods: the orientation histogram-based shape prior method of [9] (Met 1); our method using contextual information in graph cuts (GCCont), the method in [39] using object interaction priors (Met 2) and the multilabel graph cuts method of [38] (Met 3). The automatic segmentations were compared with manual segmentation using DM and HD. Shape prior methods are used for comparison to highlight the advantages of including contextual information. We also summarize the results of LV segmentation for different algorithms used in the STACOM segmentation challenge [48]. Note that [48] provides only the results of LV segmentation over the 100 patient datasets, while we provide LV (and RV) segmentation results for 30 datasets. While these set of results may not be exactly comparable, it provides a fair indication of our method's performance on the same database.

To fix λ, we adopt the following steps. We choose a small subset of the training data consisting of 10 patient volumes, and perform segmentation using our method but with λ varying from 0 to 1 in steps of 0.01. Based on the segmentation accuracy using DM, we set λ=0.05

Segmentation Results

Table 1 summarizes the performance of different algorithms for LV segmentation and Table 2 summarizes the performance of four methods for RV segmentation. In addition to the four methods mentioned above, in Table 1 we also show results for the seven methods used for the same datasets for which LV segmentation results were published in [48]. Our method gives the best performance as shown by all the different evaluation metrics. HD values (in pixels) are shown only for the first four methods because we did not have access to the segmentations generated by the other methods. Analysis of the results shows that the inclusion of context information contributes to a significant increase in segmentation accuracy, higher than when using prior

Table 1 Comparative performance of segmentation accuracy of LV (real patient images) using different methods and metrics

	DM	JI	Sen	Spe	PPV	NPV	HD
Met 1	0.87 (0.12)	0.79 (0.11)	0.91 (0.09)	0.88 (0.13)	0.85 (0.11)	0.91 (0.10)	3.2 (0.7)
GC_{Cont}	0.93 (0.06)	0.87 (0.09)	0.96 (0.0.5)	0.93 (0.09)	0.91 (0.09)	0.96 (0.06)	1.5 (0.2)
Met 2	0.91 (0.12)	0.83 (0.10)	0.94 (0.07)	0.89 (0.12)	0.89 (0.11)	0.93 (0.08)	2.2 (0.4)
Met 3	0.85 (0.13)	0.74 (0.16)	0.89 (0.15)	0.82 (0.14)	0.84 (0.15)	0.88 (0.16)	3.9 (0.9)
SCR	0.83 (0.11)	0.73 (0.14)	0.78 (0.15)	0.96 (0.04)	0.92 (0.07)	0.87 (0.08)	
INR	0.68 (0.17)	0.53 (0.17)	0.75 (0.24)	0.73 (0.16)	0.66 (0.14)	0.85 (0.12)	
AO	0.86 (0.09)	0.76 (0.12)	0.90 (0.12)	0.87 (0.09)	0.83 (0.09)	0.94 (0.06)	
DS	0.80 (0.14)	0.68 (0.16)	0.79 (0.16)	0.88 (0.08)	0.82 (0.13)	0.87 (0.09)	
EM	0.88 (0.07)	0.80 (0.10)	0.89 (0.10)	0.91 (0.08)	0.89 (0.09)	0.93 (0.06)	
AU	0.88 (0.09)	0.80 (0.13)	0.85 (0.11)	0.96 (0.04)	0.93 (0.09)	0.90 (0.08)	
NU	0.75 (0.10)	0.61 (0.12)	0.63 (0.12)	0.99 (0.02)	0.96 (0.06)	0.81 (0.06)	

Since the segmentations of other methods were not available, we could not calculate their HD values. Values indicate the mean and standard deviation

PPV positive predictive value, *NPV* negative predictive value

Table 2 Comparative performance of segmentation accuracy of RV using FOUR methods and two metrics (DM and HD). Values indicate the mean and standard deviation

	Dice metric (%)				HD (pixels)			
	Met 1	GC_{Cont}	Met 2	Met 3	Met 1	GC_{Cont}	Met 2	Met 3
RV	88.2±0.9	92.9±1.2	90.4±1.0	84.9±1.1	3.9±0.7	1.5±0.3	2.1±0.4	4.4±0.6

shape information. Although shape priors provide high level knowledge, contextual knowledge is much more informative in segmentation. This is because context information encodes higher level knowledge with respect to other regions in an image, while shape priors encode knowledge about the one shape in question.

A statistical measure of segmentation performance was determined by conducting *t* tests on the set of HD and DM values obtained by Met1, Met2, Met3, and GCCont. Comparing between Met 3 and all other methods, we get $p<0.015$ indicating a large difference in segmentation output. This is expected because Met 3 uses only intensity information while all the other methods use prior knowledge to different degrees. Comparing between Met 1 and Met 2, we obtain $p<0.04$ which means a small degree of significance between the two sets of results. This leads us to conclude that the two measures use similar level of information, but Met 2 integrates a little more relevant information because of the use of global context. However, for GCCont, we get $p<0.024$ when compared with Met 1 and Met 2 which is an indication of fairly significant improvement over the two methods.

Figure 2 shows segmentation results for different algorithms. While prior shape information is helpful, we observe that context information improves segmentation accuracy as observed by the results of GCcont and Met 2.

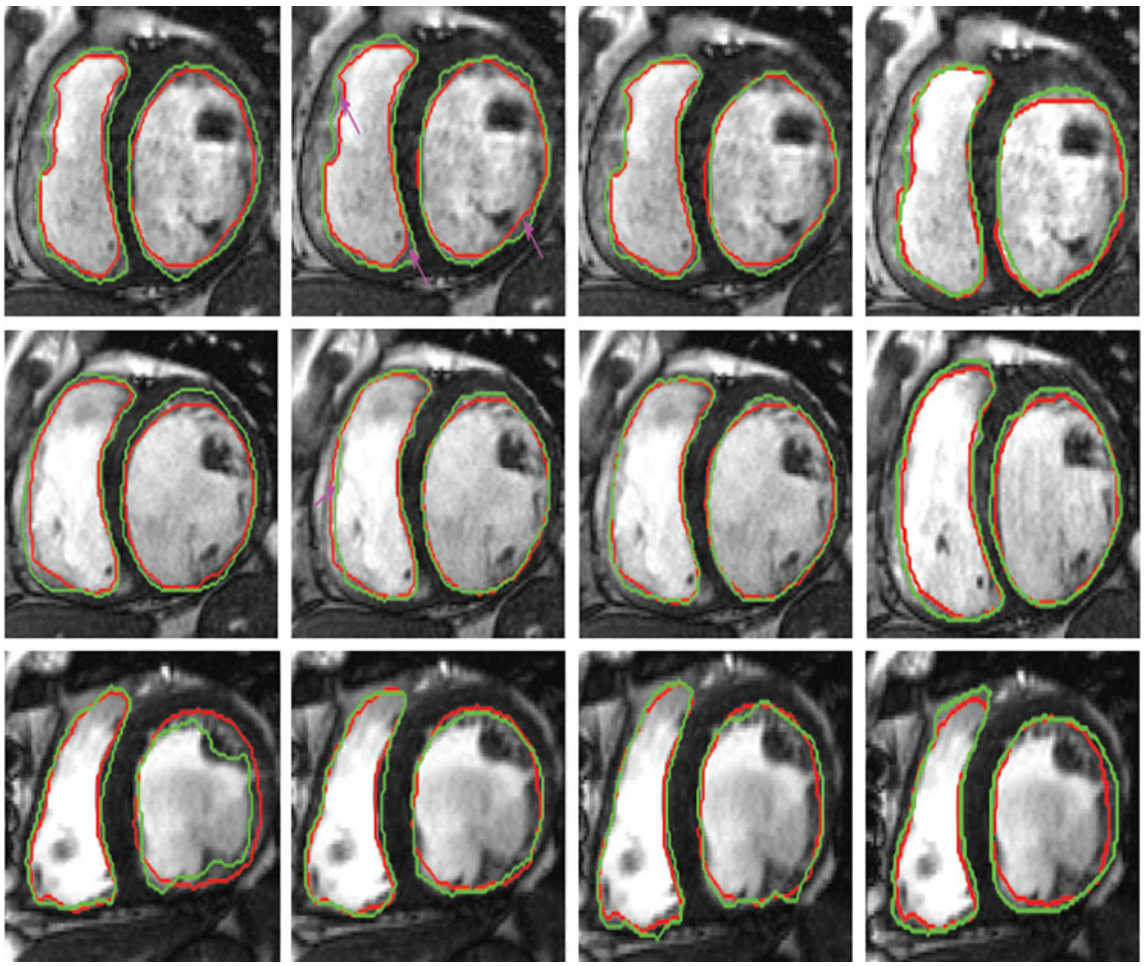

Fig. 2 Segmentation results for different methods. Each row shows results of different patient datasets from the STACOM database. *Red contour* shows the manual segmentations, and green contours show results for automatic segmentations. Columns *1–4* show segmentation outputs for Met 1, Met 2, GCcont, and Met 3

It is interesting to note that Met 1 and Met 2 have similar performance. Both context and shape priors give helpful knowledge to aid segmentation. Met 3 relies only on intensity to achieve a high level of segmentation accuracy. Information and even though it is a multilabel optimization method intensity information is not sufficient The advantage of context information is its ability to encode more global knowledge into the segmentation than shape priors. These methods are able to capture a wider range of shape changes than shape priors over the same training data. Our method compares favorably with Met 2 and in many cases gives better segmentation accuracy. Although Met 2 and GCcont are similar in terms of the geometric features used, our method is simpler as we do not need to introduce submodular graphs for graph cut optimization. Additionally, we incorporate interpixel relationships into the smoothness cost based on training data which contributes to higher segmentation accuracy. A t test between results of GCCont and Met 2 gives $p<0.03$ indicating statistically different results.

Segmentation Results for Added Noise

Figure 3 shows segmentation results after various iterations of our method under different levels of added noise. The first row corresponds to results on the original image (from the STACOM dataset) without any added noise. The first column shows the initial segmentations of the RV and LV, followed by the updated segmentations (after each iteration) in each subsequent column. The second row shows segmentation results for the image to which zero mean Gaussian noise (of variance 0.01) has been added, and the third row shows results for zero mean Gaussian noise of variance 0.05. With addition of noise, the initial segmentations tend to leak out to the background. However, the segmentations are refined at each subsequent iteration and the final output is quite close to the desired segmentations. The average DM for the initial segmentations of RV in the three cases are 85.3, 82.1, and 80.3 with the corresponding HD values being 4.2, 5.1, and 5.9 pixels. The average DM value of the final segmentations are 90.3, 88.2, and 86.7, and the corresponding HD are 1.9, 2.3, and 3.2. These values indicate a large improvement in segmentation over multiple iterations. A series of t tests on the corresponding DM and HD values after the initial and final iterations gives $p<0.025$, indicating statistically significant improvements.

For LV, the average initial DMs are 86.3, 84.2, and 81.5 and the corresponding HD values being 3.8, 4.8, and 5.3 pixels. The DM value of the final segmentations are 91.4,

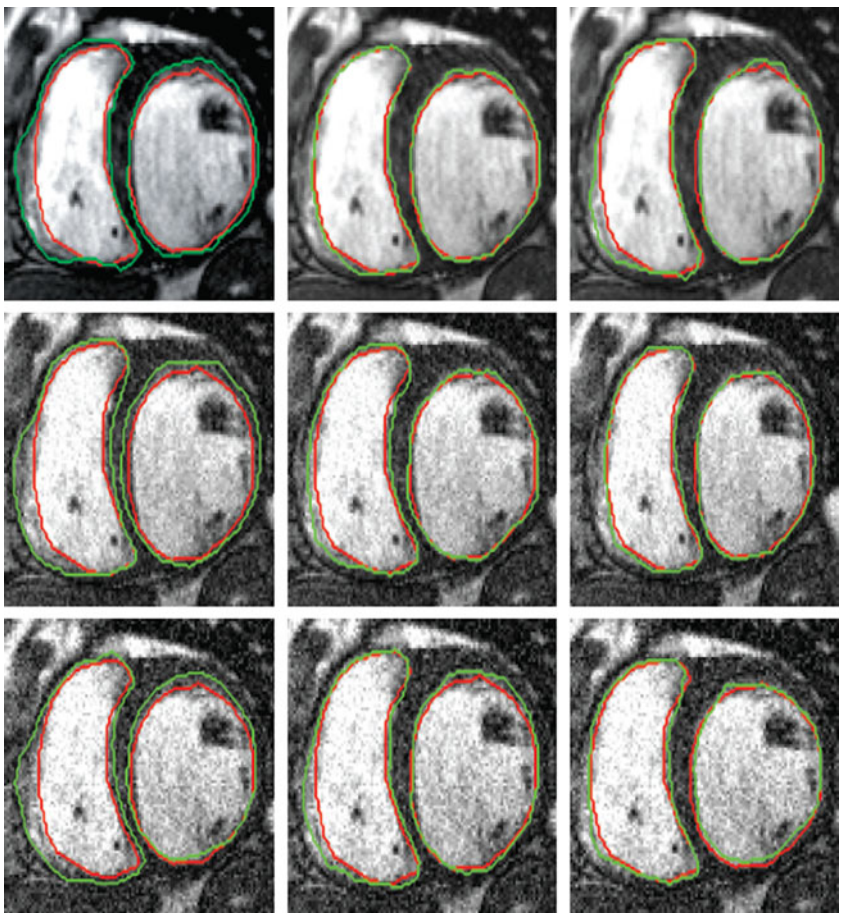

Fig. 3 Segmentation results of our method for different noise levels. The *first row* shows results for original image. The *second* and *third row* show results with added Gaussian noise of zero mean and variance of 0:01 and 0:05, respectively. *Red contour* shows the manual segmentations, and green contours show results for automatic segmentations. Columns *1–3* show segmentation outputs after first, second, and third iterations

Table 3 Performance of our method for different levels of added noise

		μ=0, σ=0.005	μ=0, σ=0.01	μ=0, σ=0.08
LV	DM	86.1±1.2	81.1±0.9	76.8±1.3
	HD	2.2±0.7	3.1±0.2	4.4±0.4
RV	DM	84.2±0.9	80.5±1.2	75.1±0.5
	HD	2.8±0.5	3.7±0.3	4.9±0.3

The image intensities were in the range 0 and 1

μ mean, σ variance of added noise

88.8, and 87.2, and the corresponding HD are 1.7, 2.1, and 2.9. *t* tests for DM and HD values after initial and final segmentation gives $p<0.033$, which also indicates significant improvement after iterations. The average DM and HD measures for different levels of noise are summarized in Table 3.

Performance for Initial Undersegmentation

In another set of experiments, we deliberately undersegment the RV and LV in the first step and apply our method. Figure 4 shows the visual results of our method for different degrees of undersegmentation. When the undersegmentations have DM >62, the final segmentations have DM ≥80. When the initial DM ≤62, the final segmentations have DM <80. In [27], the authors report that DM ≥80 indicates good agreement with manual segmentations for cardiac images. Table 4 summarizes the final segmentation results for different degrees of undersegmentation. A highly undersegmented initial image does not provide sufficient geometric information for the subsequent segmentations to approach the actual shape. The same holds true for oversegmentations. The threshold undersegmentation (or oversegmentation) degree for a good final segmentation would depend upon the degree of shape variations captured from images in the training data.

Table 4 Performance of our method for different degrees of initial undersegmentation. DM indicates the range of DM for the initial undersegmentations. Values indicate mean and standard deviation

Initial DM		79–83	70–79	61–69
LV	Final DM	91.0±1.2	84.1±0.9	78.8±1.3
	HD	1.7±0.7	2.8±0.2	4.1±0.4
RV	Final DM	89.2±0.9	82.5±1.2	77.1±0.5
	HD	1.9±0.5	2.7±0.3	4.2±0.3

An additional factor influencing the accuracy of the final segmentation is the search range or the width of the search band in each iteration. A low search range will not be able to recover the actual shape from a under/oversegmented initial shape. We set a search width of ±20 pixels after the following considerations. We had made available the manual segmentations for different phases of the cardiac cycle where the size of the RV and LV are changing. For one dataset, we found the mean of the segmentations and the difference contour from the mean for each phase. The maximum distance between contours for a single patient was found to be 18 pixels. Therefore, we set the threshold distance as 20 pixels in each direction.

Conclusion

We have proposed a novel graph cut framework for cardiac MR image segmentation that incorporates context information

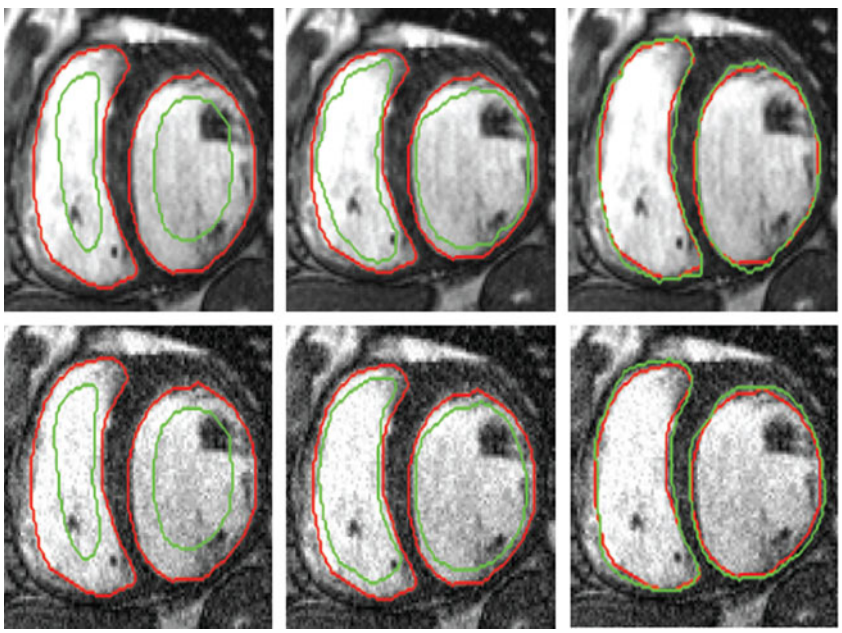

Fig. 4 Segmentation results of our method for different degrees of undersegmentation. The *first* and *second* rows show results for an initial undersegmentation of DM= 70 % and DM=74 %. *Red contour* shows the manual segmentations and *green contours* show results for automatic segmentations. Columns *1–3* show segmentation outputs after first, second, and third iterations

from RV and LV for improved segmentation accuracy of each. Context information was modeled in the form of relative orientation of the two organs from a set of training images in which the LV and RV were manually segmented. The learned orientation relationship is used to formulate two penalty and smoothness costs that lead to improved segmentation performance. Our formulation avoids the need to construct submodular graphs as given in [39]. Experimental results on publicly available patient datasets show the efficacy of our method compared to previous shape prior-based methods. Although shape priors also incorporate knowledge, contextual information provides a more global knowledge from the same training dataset and hence improves segmentation. Tests on experiments simulating different levels of noise and inaccurate initial segmentations show that our method is able to recover the desired final output in three to four iterations of graph cut optimization.

References

1. Allender S: European Cardiovascular Disease Statistics. European Heart Network. 2008
2. A.F. Frangi, W.J. Niessen, and M.A. Viergever, "Three dimensional modeling for functional analysis of cardiac images: a review," IEEE Trans Med. Imag, vol. 20, no. 1, pp. 2–25, 2001
3. C. Petitjean and J-N. Dacher, "A review of segmentation methods in short axis cardiac mr images," Med. Imag. Anal., vol. 15, no. 2, pp. 169–184, 2011
4. S. Shors, C. Fung, C. Francois, P. Finn, and D. Fieno, "Accurate quantification of right ventricular mass at MR imaging by using cine true fast imaging with steady state precession: study in dogs.," Radiology, vol. 230, no. 2, pp. 383–388, 2004
5. Jolly MP: Automatic recovery of the left ventricle blood pool in cardiac cine MR images. In: MICCAI, 2008, pp 110–118
6. N. Paragios, "A variational approach for the segmentation of the left ventricle in cardiac image analysis," Intl. J. Comp. Vis., vol. 50, no. 3, pp. 345–362, 2002
7. M. Lynch, O. Ghita, and P. Whelan, "Left ventricle myocardium segmentation using a coupled level set with a-priori knowledge," Comput. Med. Imag. Graph., vol. 30, no. 4, pp. 255–262, 2006
8. Lin X, Cowan B, and Young A: Model based graph cut method for segmentation of the left ventricle. In In Proc: EMBC, 2005, pp 3059–3062
9. Mahapatra D and Sun Y: Orientation histograms as shape priors for left ventricle segmentation using graph cuts. In Proc: MICCAI, 2011, pp 420–427
10. Mahapatra D: Cardiac image segmentation from cine cardiac MRI using graph cuts and shape priors. Journal of Digital Imaging. doi:10.1007/s10278-012-9548-5, 2013
11. J. Cousty, L. Najman, M. Couprie, S. Clment-Guinaudeau, T. Goissen, and J. Garot, "Segmentation of 4-D cardiac MRI: automated method based on spatio temporal watershed cuts.," Image and Vis. Comput., vol. 28, no. 8, pp. 1229–1243, 2010
12. C. Cocosco, W. Niessen, T. Netsch, E-J. Vonken, G. Lund, A. Stork, and M. Viergever.,"Automatic image driven segmentation of the ventricles in cardiac cine MRI.," J. magn. Reson. Imag., vol. 28, no. 2, pp. 366–374, 2008
13. Mahapatra D, and Sun Y: Joint registration and segmentation of dynamic cardiac perfusion images using mrfs. In In Proc: MICCAI, 2010, pp 493–501
14. D. Mahapatra and Y. Sun, "Integrating segmentation information for improved elastic registration of perfusion images using an mrf framework," IEEE Trans. Imag. Proc., vol. 21, no. 1, pp. 170–183, 2012
15. Mahapatra D: Joint segmentation and groupwise registration of cardiac perfusion images using temporal information. Journal of Digital Imaging
16. D. Mahapatra: Groupwise registration of dynamic cardiac perfusion images using temporal dynamics and segmentation information", SPIE Medical Imaging 2012, SPIE Vol 8314, pp 1–7
17. M.R. Kaus, J. von Berg, J. Weese, W. Niessen, and V. Pekar, "Automated segmentation of the left ventricle in cardiac MRI," Med Image Anal., vol. 8, no. 3, pp. 245–254, 2004
18. Zhu Y, Papademetris X, Sinusas A, and Duncan J.S: Segmentation of left ventricle from 3d cardiac mr image sequence using a subject specific dynamic model. In Proc.IEEE CVPR, 2008, pp 1–8
19. Sun W, Setin M, Chan R, Reddy V, Holmvang G, Ch V, and Willsky A: segmenting and tracking of the left ventricle by learning the dynamics in cardiac images. In Proc. IPMI, 2005, pp 553–565
20. R. H. Davies, C. J. Twining, T. F. Cootes, J. C. Waterton, and C. J. Taylor, "A minimum description length approach to statistical shape modelling," IEEE Trans. Med. Imag., vol. 21, pp. 525–537, 2002
21. Perperidis D, Mohiaddin R, and Rueckert D: Construction of a 4d statistical atlas of the cardiac anatomy and its use in classification. In MICCAI, 2005, pp 402–410
22. Besbes A, Komodakis N, and Paragios N: Graph-based knowledge-driven discrete segmentation of the left ventricle. In IEEE ISBI, 2009, pp 49–52
23. S.C. Mitchell, B.P.F. Lelieveldt, R.J. van der Geest, H.G. Bosch, J.H.C Reiver, and M. Sonka, "Multistage hybrid active appearance models: segmentation of cardiac MR and ultrasound images," IEEE Trans Med. Imag, vol. 20, no. 5, pp. 415–423, 2001
24. H. Zhang, A. Wahle, R. Johnson, T. Scholz, and M. Sonka, "4-D cardiac MR image analysis: left and right ventricular morphology and function.," IEEE Trans Med. Imag, vol. 29, no. 2, pp. 350–364, 2010
25. Zambal S, Hladuvka J, and Buhler K: Improving segmentation of the left ventricle using a two component statistical model. In MICCAI, 2006, pp 151–158
26. Lelieveldt B, Mitchell S, Bosch J, van der Geest R, Sonka M, and Reiber J: Time continuous segmentation of cardiac image sequences using active appearance motion models. In IPMI, 2001, pp 446–452
27. C. Pluempitiwiriyawej, J.M.F. Moura, Y.L.Wu, and C. Ho, "STACS: new active contour scheme for cardiac MR image segmentation," IEEE Trans. Med. Imag., vol. 24, no. 5,pp. 593–603, 2005
28. Billet F, Sermeanst M, Delingette H, and Ayache N: Cardiac motion recovery and boundary conditions estimation by coupling an electromechanical model and cine-MRI data. In Functional imaging and modeling of the heart (FMIH), 2009, pp 376–385
29. J. Ltjnen, S. Kivist, J. Koikkalainen, D. Smutek, and K. Lauerma, "Statistical shape model of atria, ventricles and epicardium from short- and long-axis MR images," Med Image Anal., vol. 8, no. 3, pp. 371–386, 2004
30. S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images.," IEEE Trans. Patt. Anal. Mach. Intell., vol. 6, no. 6, pp. 721–741, 1984
31. Kumar S and Hebert M: Discriminative random fields: a discriminative framework for contextual interaction in classification. In Proc. ICCV, 2003, pp 1150–1157

32. S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," IEEE Trans. Patt. Anal. Mach. Intell., vol. 24, no. 24, pp. 509–522, 2002
33. Hoiem D, Efros AA, and Hebert M: Putting objects in perspective. In Proc. CVPR, 2006, pp 2137–2144
34. He X, Zemel RS, and Carreira-Perpinan MA: Multiscale conditional random fields for image labeling. In Proc. CVPR, 2004, pp 695–702
35. Murphy K, Torralba A and Freeman WT: Graphical model for recognizing scenes and objects. In Proc. NIPS
36. Z. Tu and X. Bai, "Auto-context and its application to high-level vision tasks and 3d brain image segmentation," IEEE Trans. Patt. Anal. Mach. Intell., vol. 32, no. 10, pp. 1744 – 1757, 2010
37. Li W, Liao S, Feng Q, Chen W, and Shen D: Learning image context for segmentation of prostate in ct-guided radiotherapy. In MICCAI, 2011, pp 570–578
38. Delong A and Boykov Y: Globally optimal segmentation of multi-region objects. In ICCV, 2009, pp 285–292
39. Ben Ayed I, Punithakumar K, Garvin G, Romano W, and Li S: Graph cuts withinvariant object-interaction priors: Application to intervertebral disc segmentation. In IPMI, 2011, pp 221–232
40. Song Q, Chen M, Bai J, Sonka M, and Wu X: Surface-region context in optimal multi-object graph based segmentation: robust delineation of pulmonary tumors. In IPMI, 2011, pp 61–72
41. Y. Boykov and O. Veksler, "Fast approximate energy minimization via graph cuts," IEEE Trans. Pattern Anal. Mach. Intell., vol. 23, pp. 1222–1239, 2001
42. V. Chalana and Y. Kim, "A methodology for evaluation of boundary detection algorithms on medical images," IEEE Trans. Med. Imag., vol. 16, no. 5, pp. 642–652, 1997
43. D.P. Huttenlocher, G.A. Klanderman, and W.J. Rucklidge, "Comparing images using the hausdorff distance," IEEE Trans. Pattern Anal. Machine Intell., vol. 15, no. 9, pp. 850–863, 1993
44. C.G. Fonseca, M. Backhaus, D.A. Bluemke, R.D. Britten, J.D. Chung, B.R. Cowan, I.D. Dinov, J.P. Finn, P.J. Hunter, A.H. Kadish, D.C. Lee, J.A.C. Lima, P. Medrano-Gracia, K. Shivkumar, A. Suinesiaputra, W. Tao, and A.A. Young., "The cardiac atlas project an imaging database for computational modeling and statistical atlases of the heart.," Bioinformatics, vol. 27, no. 16, pp. 2288–2295, 2011
45. A.H. Kadish, D. Bello, J.P. Finn, R.O. Bonow, A. Schaechter, H. Subacius, C. Albert, J.P. Daubert, C.G. Fonseca, and J.J. Goldberger., "Rationale and design for the defribrillators to reduce risk by magnetic resonance imaging evaluation (determine) trial.," J. Cardiovascular Electrophysiology, vol. 20, no. 9, pp. 982–987, 2009
46. A. Elen, J. Hermans, J. Ganame, D. Loeckx, J. Bogaert, F. Maes, and P. Suetens., "Automatic 3-d breath-hold related motion correction of dynamic multislice mri.," IEEE Trans. Med. Imag., vol. 29, no. 3, pp. 868–878, 2010
47. A.A. Young, B.R. Cowan, S.F. Thrupp, W.J. Hedley, and L.J. DellItalia., "Left ventricular mass and volume: fast calculation with guide-point modeling on mr images.," Radiology, vol. 202, no. 2, pp. 597–602, 2000
48. Suinesiaputra A and et al: Left ventricular segmentation challenge from cardiac mri: a collation study. In STACOM 2011, 2011, pp 88–97