Segmentation of RV in 4D Cardiac MR Volumes using Region-Merging Graph Cuts

Oskar M O Maier[1], Daniel Jiménez[1], Andrés Santos[1,2], María J Ledesma-Carbayo[1,2]

[1] ETSI de Telecomunicación, Universidad Politécnica de Madrid, Spain
[2] CIBER de Bioingeniera, Biomateriales y Nanomedicina (CIBER-BBN), Spain

Abstract

Non-invasive quantitative assessment of the right ventricular anatomical and functional parameters is a challenging task. We present a semi-automatic approach for right ventricle (RV) segmentation from 4D MR images in two variants, which differ in the amount of user interaction. The method consists of three main phases: First, foreground and background markers are generated from the user input. Next, an over-segmented region image is obtained applying a watershed transform. Finally, these regions are merged using 4D graph-cuts with an intensity based boundary term. For the first variant the user outlines the inside of the RV wall in a few end-diastole slices, for the second two marker pixels serve as starting point for a statistical atlas application. Results were obtained by blind evaluation on 16 testing 4D MR volumes. They prove our method to be robust against markers location and place it favourably in the ranks of existing approaches.

1. Introduction

Cardiovascular disease is the number one cause of death in the western world. In the management of most cardiac disorders, an evaluation of the myocardial function and structure is important. Most cardiopulmonary diseases require a proper assessment of the right ventricle (RV) chamber, a challenging task given its shape, myocardial thickness and the intracavity structure.

Magnetic resonance (MR) is the de-facto standard in cardiac assessment and derived parameters correlate well with the real world.

So far, the standard approach in clinical praxis is manual segmentation in the end-diastole (ED) and end-systole (ES) phases[1], with reported times between 9.8 and 20 minutes depending on the observers experience. Furthermore [1] show that expert knowledge is required to achieve a reliable and reproducible segmentation.

[1] Note that we refer to steps through time as *phases* and through space in axial view as *slices*.

While numerous segmentation algorithms for the left ventricle (LV) have been published, the RV has been seldom approached. Furthermore the majority of the methods are little general and the reported results not comparable. The recent overview article [2] summarizes the main problems and approaches and shows that the best performing methods rely heavily on prior knowledge.

RV segmentation from MR is considered one of the most challenging task in organ segmentation. Grosgeorge et al. [3] identify as the three main challenges *a)* the fuzziness of the cavity borders due to blood flow, acquisition artefacts, and partial volume effects, *b)* the trabeculations (i.e. wall irregularities) and *c)* the complex half-moon shape that varies strongly over the heart cycle. We would like to extend this list by *d)* low inter-slice correspondence in axial view due to large slice-thickness ($7-12$mm) and *e)* breathing artefacts.

We propose a novel approach to semi-automatically segment the RV endocardium surface based on 4D region-merging graph cuts (rmGC). Two variants are presented: One with medium user-interaction and one using statistical atlases. Additionally we contribute with an automatic method to enhance the manual markers required by the graph cut (GC) algorithm.

The following section is dedicated to the presentation of our method. Experiments and evaluation results can be found in Sec. 3. Sec. 4 concludes and proposes several future lines.

2. Method

Our approach consists of various steps detailed in the schematic overview of Fig. 1. The most important components are described in this section.

2.1. Region-merging graph cut

Graph cuts are a popular and powerful tool in image processing. Based on graph theory, they minimize an energy function that can be formulated in various ways [4]. This flexibility has earned them many applications in organ seg-

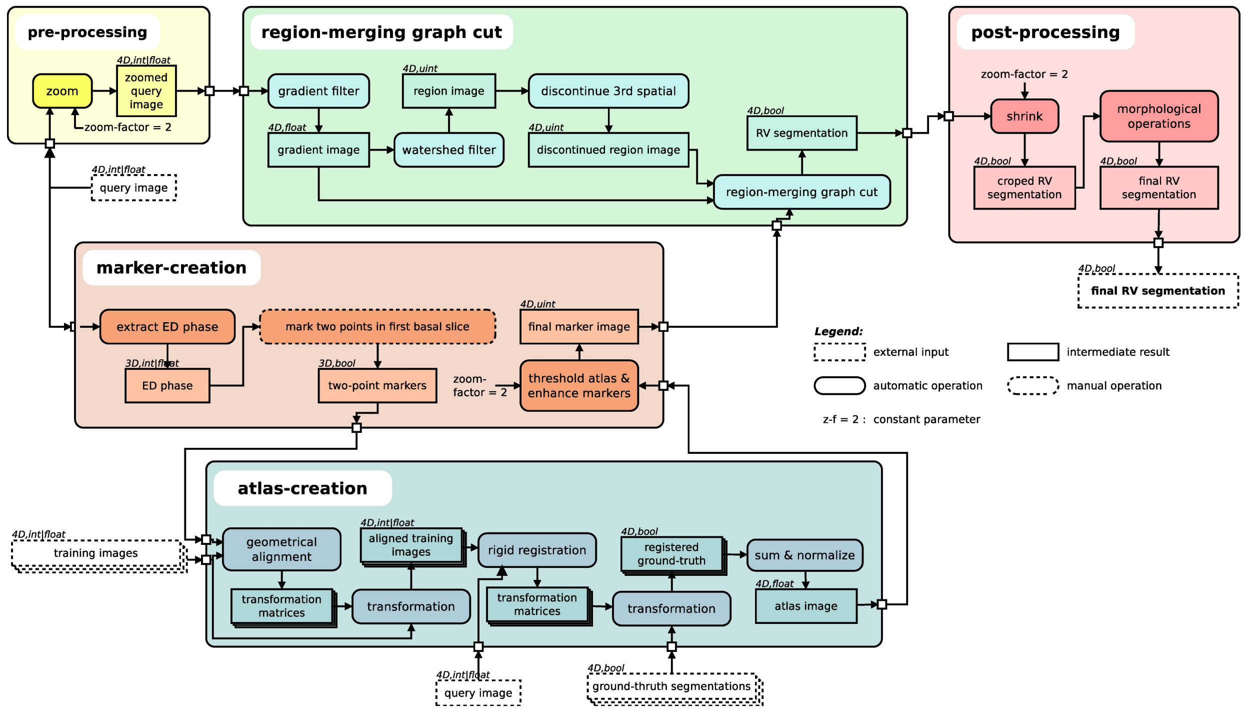

Figure 1. Schema of the proposed approach, variant II. For variant I the atlas creation group is removed.

mentation in general [5, 6] and also ventricle segmentation [7]. To our best knowledge, only [5] ever attempted an application of GCs to the RV and only in 2D slices.

Although an efficient implementation is available, the classical voxel-based GC is very memory consuming. For large or multi-dimensional images, such as often processed in the medical domain, they are therefore less suitable. This is also reflected in the state of the art, with only very rare applications to 4D volumes.

The most popular propositions to reduce the graph size of GCs are watersheds (WS) (e.g. [8]). The idea consists of applying a WS filter as a pre-processing step and subsequently joining the resulting regions together using GCs.

The WS filter segments an image into many regions and is usually applied to the gradient magnitude map. Over-segmentation and edge-preservation render it an ideal rmGC pre-processing step.

For the graph cut energy term formulation we follow the boundary term proposed by [6, 8] adapted to 4D. It takes intensity differences along the regions borders as well as the length of the shared wall into account. No regional term was employed.

2.2. Variation I: Medium user interaction

The first evaluated approach requires the user to outline the midline of the RV wall in 4 to 5 slice of the ED phase. Via dilation and/or erosion we can extract foreground- (FG) and background- (BG) markers from these simple contours. In this work we only evaluate the endocardium segmentation, but the approach is equally applicable for the epicardium.

2.3. Variation II: Minimal user interaction

To minimise the user interaction, we propose a second method using registration and statistical atlases. Based on solely two points marked in the first basal slice of the ED phase, we initially align training images to the query image. Then we obtain the transformation matrices using a single-resolution rigid registration with a stochastic gradient descent optimizer and the advanced mattes mutual information metric in *elastix* [9] in the ED and ES phases. These matrices are used to rotate and shift the ground-truth segmentations of the training images, which are subsequently combined into a statistical atlas of the RV. Finally the FG- and BG-markers for the endocardium are determined via thresholding. Varying the thresholds, also in this variant the epicardium markers could be extracted.

2.4. Marker enhancement

Since a stronger marker-base makes the GC perform more robust, we enhance these sparse markers at ED (variant I) resp. ED and ES (variant II) phases automatically. As first measure we copy the ED markers to the opposite phase of the 4D volume, which is again ED. Furthermore we propagate the FG markers starting from both ED phases towards the ES phase, applying erosion at each step.

3. Experiments

3.1. Set-up

Our results were obtained on the dataset published in the scope of the MICCAI 2012 RV-Challenge [1, 10], consisting of 16 training cases with ground-truth segmentation and 16 testing images without. Evaluations on the testing data was performed by the workshop organisers acting as independent referee and is therefore guaranteed to be unbiased. Computed measures include the dice metric (DM), the Hausdorf Distance (HD) and clinical parameters.

The manual markers for both variants were created by two observers which each delivered two sets of markers. Thus the inter- as well as intra-observer variability can be assessed. For variant I 2-3 minutes/volume were required, for variant II a mere 2-3 seconds/volume.

3.2. Evaluation

3.2.1. Variation I

Results were obtained on the testing set for all marker sets and are listed in Table 1.

High DM scores were obtained, especially in the ED where the manual markers are situated. The HDs are relatively high, caused by peak leakages observable during visual examination. The intra- and inter-observer difference are only minimal, as the standard deviations show clearly. Our proposed method is able to follow the RV from ED to ES phase in most cases and is very robust against the placement of the initial markers.

Clinical parameters are depicted in Table 2, the average EF mean error is 0.1963.

Table 2. V-I: Clinical parameters obtained on the testing set for the ED and ES phases and their total mean. R=correlation coefficient (ideal=1), *a* and *b*=coefficients of a linear regression fit ($y = ax + b$, ideally *a*=1 and *b*=0).

	Average (std) over markersets		
	R	a	b
Endo area	$0.97(0.00)$	$1.00(0.01)$	$1.50(0.51)$
ED volume	$0.99(0.00)$	$1.06(0.05)$	$1.02(2.34)$
ES volume	$0.96(0.02)$	$1.06(0.02)$	$6.73(2.50)$
EF	$0.86(0.02)$	$1.02(0.07)$	$-0.06(0.03)$

Again our method shows little intra- and inter-observer variability. The obtained values for R and a are near their optimum. This is not true for the b coefficient, which has sometimes very high values. This means that our method has a tendency to overestimate the RV volume that is especially noticeable in the ES phase.

Following [3] we also assessed the DM and HD separately for three slice (base, mid, apex). The results are presented in Table 3. In ED and ES phase, the best results are obtained in the basal slices and the worst in the apical slices. This is consistent with the observations made by [3]. Compared with their evaluation on the RV, our average DM values are higher on all slice levels and in both phases.

Table 3. V-I: Averaged (over all observers) dice metric (DM) obtained in basal, mid-cavity and apical slices. Comparison results were obtained on a different database.

			Avg. (avg. std)	
			Our method (var. I)	Grosgeorge [3]
ED	base	DM	$0,92(\pm 0,05)$	$0.80(\pm 0.01)$
	mid	DM	$0,88(\pm 0,07)$	$0.71(\pm 0.00)$
	apex	DM	$0,72(\pm 0,18)$	$0.46(\pm 0.00)$
ES	base	DM	$0,83(\pm 0,16)$	$0.59(\pm 0.01)$
	mid	DM	$0,80(\pm 0,15)$	$0.55(\pm 0.02)$
	apex	DM	$0,46(\pm 0,26)$	$0.25(\pm 0.04)$

The statistical evaluation places our approach favourably in the ranks of the existing methods. But a direct comparison is often not possible, as the results were obtained on different datasets, for different phases, for different slice levels and also using different evaluation measures.

Our algorithm runs $95.66(\pm 14.29)$ sec on average, 48.80 of which are contributed to the graph cut. Compared to standard voxel-based GC, the computation time is reduced by the factor 8 and the peak memory consumption to 18% without any loss in precision. Using the same markers, the 4D approach outperforms 3D cuts in ED and ES and is only slightly under the results of 2D+time cuts, while requiring not even half of its manual markers.

Thus the additional information encoded in a 4D volume provides advantages over the use of 3D that goes further than the fact that less manual markers are required. The rmGC's computational requirements enable the exploitation of these resources.

3.2.2. Variation II

Results were obtained on the testing dataset using the training set for the atlases and are listed in Table 4.

Table 4. V-II: Dice metric (DM, from 0 to 1) and Hausdorf Distance (HD, in mm) results on the testing set in ED phase, ES phase and their total mean.

	DM (std)	HD (std)
ED	0.78 (0.17)	14.15 (7.72)
ES	0.64 (0.25)	14.80 (6.92)
Total	0.72 (0.22)	14.44 (7.36)

As the marker points are only required for initial im-

Table 1. V-I: Dice metric (DM, from 0 to 1) and Hausdorf Distance (HD, in mm) results on the testing set in ED phase, ES phase and their total mean.

		Observer I		Observer II		Average (std)
		Markerset 1	Markerset 2	Markerset 1	Markerset 2	
ED	DM/HD	0.83 / 9.41	0.84 / 9.09	0.85 / 9.55	0.86 / 8.8	$0.84(0.01)$ / $9.21(0.29)$
ES	DM/HD	0.67 / 14.98	0.68 / 15.28	0.69 / 14.26	0.72 / 14.49	$0.69(0.02)$ / $14.75(0.40)$
Total	DM/HD	0.76 / 11.81	0.77 / 11.72	0.78 / 11.56	0.80 / 11.15	$0.780.01)$ / $11.56(0.25)$

age alignment prior to registration, no difference has been observed between the marker sets. Clearly the obtained score falls behind variant I. This is easily explained by the reduced manual interaction: In the ES phase, were for both variants no manual markers exist, the DM results do not show significant differences. The more relaxed marker placement leads to some leakages and so higher HDs.

Expectedly the ejection fraction results fall with $R = 0.75$, $a = 0.54$ and $b = 0.10$ behind variant I results. The mean error is with 0.36 (std 0.24) to high for a clinical application. A segmentation into basal, mid and apical results leads to the same observation already made for variant I, namely that the best results can be observed in the basal slices.

4. Conclusion

We proposed a semi-automatic 4D rmGC algorithm with automatic markers enhancement and applied it successfully to a large cardiac MR database. The low computational requirements render the rmGC ideal for an application to large 4D volumes, giving access to the additional information encoded in these.

The experiment proved our method to be very robust against the initial placement of the markers and it thus can be considered reproducible. An evaluation on a large dataset with medically approved ground truth placed our approach favourably in the ranks of the existing methods. Additionally various geometrical and clinical parameters have been assessed. Extracting the GC markers from a statistical atlas proved applicable, although with a negative impact on the evaluation score.

Nevertheless there remains room for improvement. Especially in the apical slices the gradient information is not sufficient for the GC to reliably segment the RV. The next step would therefore be to implement some prior information, such as shape models or statistical atlases, into the GC's energy function. Furthermore, it would be desirable to extend the method to also segment the epicardium surface. This would also allow the usage of wall-thickness constraints to avoid strong leakages, from which our method suffers in some cases. Additionally it would be desirable to develop an auto-initialization method to remove the need of manual marker placement.

Acknowledgement

This work was partially supported by Spains Ministry of Science and Innovation through CDTI CENIT (AMIT) and projects TEC2010-21619-C04-03 & TEC2011-28972-C02-02; Comunidad de Madrid (ARTEMIS S2009/DPI-1802), and the European Funds (FEDER).

References

[1] Caudron J, et al. Cardiac MRI assessment of right ventricular function in acquired heart disease: Factors of variability. Academic Radiology 2012;19(8):991–1002.

[2] Kang D, et al. Heart chambers and whole heart segmentation techniques: review. J Electr Imaging 2012; 21(1):010901–1–010901–16.

[3] Grosgeorge D, et al. Automatic cardiac ventricle segmentation in MR images: a validation study. Int J of Comput Assisted Radiol and Surg 2011;6(5):573–581.

[4] Kolmogorov V, Zabin R. What energy functions can be minimized via graph cuts? IEEE Trans Pattern Anal Mach Intell 2004;26(2):147–159.

[5] Boykov Y, et al. Segmentation of dynamic N-D data sets via graph cuts using markov models. In MICCAI'01, volume 2208 of Lect Notes in Comput Science. 2001; 1058–1066.

[6] Stawiaski J, et al. Interactive liver tumor segmentation using graph-cuts and watershed. In Grand Challenge Liver Tumor Segmentation (MICCAI'08 Workshop) 2008; .

[7] Jolly M. Automatic segmentation of the left ventricle in cardiac MR and CT images. Int J of Comput Vis 2006; 70(2):151–163.

[8] Stawiaski J, Decencire E. Region merging via graph-cuts. Image Anal and Stereology 2008;27(1):39–45.

[9] Klein S, et al. Elastix: A toolbox for intensity-based medical image registration. IEEE Trans Med Imaging 2010; 29(1):196–205.

[10] Caudron, et al. Diagnostic accuracy and variability of three semi-quantitative methods for assessing right ventricular systolic function from cardiac MRI in patients with acquired heart disease. Europ Radiol 2011;21(10):2111–2120.

Address for correspondence:

Oskar Maier

ETSI de Telecomunicación, Avda. Complutense 30, Madrid (España)

oskar.maier@googlemail.com