4D Morphological segmentation and the miccai LV-segmentation grand challenge

László Marák, Jean Cousty, Laurent Najman, and Hugues Talbot

Université Paris-Est, Laboratoire d'Informatique Gaspard-Monge, A3SI, ESIEE
{l.marak, j.cousty, l.najman, h.talbot}@esiee.fr

1 Introduction

The goal of the *Cardiac MR Left Ventricle Segmentation Challenge* at MICCAI 2009 is to compare state-of-the-art LV segmentation methods. This goal is facilitated through an evaluation system and a database of cardiac cine MR images, as well as expert contours, now freely available on the internet for research purposes. This challenge is important because the analysis and the segmentation of 3D+t sequences of MR cardiac images is fastidious, time consuming and error-prone for human operators, due to the large amount of data. Conversely, automated segmentation of cardiac images is well-known to be a challenging task.

This paper describes the method that our group submitted for the challenge. Section 2 briefly recalls the cardiac segmentation approach, based on 4D discrete mathematical morphology, introduced in [1]. This method leads to discrete segmentations, *i.e.*, binary masks (bitmaps) of the left ventricular myocardiums that are both spatially and temporally consistent. In section 3, we describe the data provided by the organizers of the challenge and the method they use to assess all segmentations. Finally, section 4 describes the necessary adaptation to our method in order to fulfill the requirements of the challenge, and presents the results obtained on the provided dataset.

2 Previous work: A morphological segmentation scheme

This section gives a quick overview of the method proposed in [1] to analyze 4D MR images of the left ventricular (LV) myocardiums of patients with recent acute myocardial infarction.

2.1 Image acquisition

The patients were examined on a 1,5 T MR scanner (Magneton Symphony®, Siemens) and dynamic, breath-held, ECG gated, cine-MR images were acquired perpendicularly to the long axis of the left ventricle from base to apex. Typical imaging parameters were 6 mm slice thickness, 1.7mm square pixels and 30-40 ms temporal resolution (more details in [2]). These acquisition parameters are similar to those of the challenge database.

For each patient, the cine MRI dataset consisted of a succession of contiguous gap-less LV short-axis 2D planes that were successively imaged over time (2D+t). The sequences were registered to the heart-cycle, and could be stacked in order to construct 3D sequences. The most basal slice included in the analysis was located just above the mitral valve within the LV cavity. To be included, the basal myocardium had to be visible in the entire circumference at end-systole. The most apical slice was chosen as the one with the smallest visible LV cavity at end-systole. Taken together, these different planes from base to apex constitute a 3D representation of the LV. The succession of these, over time, is a 3D+t representation of the LV. The images were oversampled in order to provide isotropic voxels. For each 3D+t sequence, a single mouse click on the center of the LV cavity at end-systolic time was recorded, and the images were cropped centered on the corresponding location. When a misalignment of the different slices of a same volume was observed, translation-only registration was applied.

2.2 Morphological segmentation

Our method is different from those based on deformable models that are commonly presented in the literature. Any segmentation method performs two tasks: recognizing the objects of interest and delineating their contours. Whereas model-based methods often perform these two tasks at the same time by minimizing some energy including internal forces (recognition) and external forces (delineation), the classical scheme in mathematical morphology (MM) separates these two tasks. More precisely, segmentation schemes in MM [3] comprise, in general, three main steps: recognition, delineation and smoothing, performed sequentially. Recognition is the process of determining the rough whereabouts of the objects. Delineation consists of the precise spatial localization of the objects borders. Finally, smoothing can be defined as the process of matching the smoothness properties of the segmented object with the a priori smoothness properties of the ground truth.

Many operators in MM consists of analyzing (unknown) geometrical objects (or more generally grayscale images) through their interaction with predefined shapes, called structuring elements. These operators can, in particular, select sets of pixels based on their shape, contrast or topological properties. They are well adapted to the recognition task and can be used to extract sets of points, called markers, that roughly correspond to the objects of interest. The sequence of operators chosen for recognition constitutes the knowledge-based part of the morphological segmentation. It must be adapted to each particular application by taking into account properties on the objects to be segmented. The delineation step is, in general, devolved to the watershed [4–6]. This operator looks, based only on the image contrast, for the "best" contours between the recognized markers. Finally, the smoothing step, which can also be done thanks to MM operators, filters the delineated objects by removing their non-significant parts, with respect to prior knowledge.

Based on this framework, we presented a new automated method [1] to segment both the endocardial and epicardial borders in 4D (3D+t) cine-MR

images. The endocardial border is segmented using a geodesic reconstruction - a morphological region growing technique - of a marker in a set of voxels detected as potential candidates. In order to recognize interior and exterior markers of the epicardial border, we use i) an exact Euclidean distance transforms [7] to take into account prior geometric properties and ii) homotopic transforms [8] to guarantee topological soundness. The delineation process is then devolved to watershed cuts [5, 6]. This operator takes as input the 4D graph -each voxel is adjacent to its 6 neighbors in 3D plus the voxels just after and before in the sequence -associated to the 3D+t sequence and a weighting function that assigns a gradient (either spatial or temporal) value to each pair of adjacent voxels. From these data, the watershed cut optimally separates the marker obtained from the recognition step, using the minimum spanning tree weights as criterion. Finally, the smoothing step is computed through sequences of morphological openings and closings. From these binary masks, we extracted the 3D surfaces of the LV myocardium at each time step – thanks to a marching cube algorithm. Thus, we obtain a succession of surfaces over time that constitutes a 4D segmentation, which is both spatially and temporally consistent.

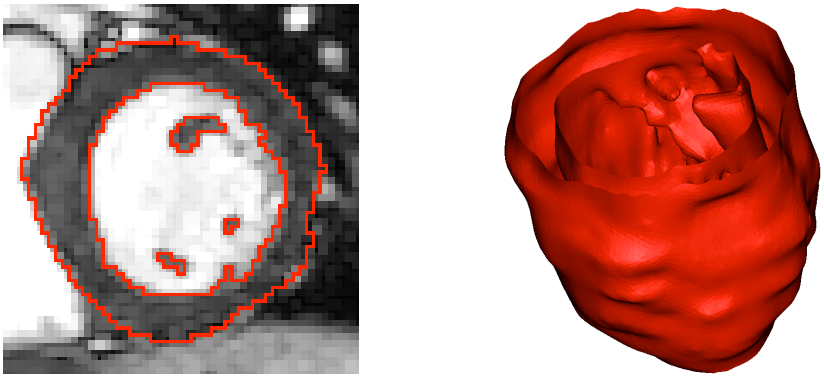

Fig. 1. Example of segmentation produced by our method. Left: the LV myocardium contours are superimposed in red. Right: 3D rendering of the segmentation.

2.3 Evaluation

This method was evaluated on cine-MR image sequences of 18 patients. Through experiments performed on this database, we demonstrated the following strong points:

– A good accuracy of the automated method compared to manual segmentations performed by two cardiologists. The mean distances between manual and automated segmentations were 1.5mm ± 0.38 and 1.8mm ± 0.38 for respectively the endocardial and epicardial borders.

– The ability of the method to compute reliable characteristics of the LV: ejection fraction (EF) and left ventricular mass (LVM). Over all the patients of the database, the automated method achieved a mean deviation on the EF (resp. LVM) of 4% (resp. 7%): a value comparable to the inter-expert deviation;
– The temporal continuity of the resulting automated segmentation. In particular, we have quantitatively shown that temporal consistency is better preserved when using one 4D watershed for the whole sequence compared to one 3D watershed per time step.
– The time-efficiency (about 3mn to segment a sequence of 25 3D-images on a low-end computer) of the proposed method; and
– the robustness of the few parameters whose settings rely mostly on physical and anatomical facts.

Furthermore, in an effort to promote open science, the database (see also [2]) used for validation was made freely available on the internet. For each one of the 18 patients, it contains 3D+t cine-MR images of the LV, together with three associated segmentations: two hand made segmentations - each one of them performed by an independent and blinded expert cardiologist - and one 4D automated segmentation obtained by the method described above. The web-address of this database is http://laurentnajman.org/heart.

3 MICCAI Challenge

3.1 Data: images and manual segmentation

Images from the MICCAI challenge contain slices that do not fit the guideline described in Section 2.1. However, our software includes an interactive graphical interface that allows the user to select the slices that follow the guideline. Our method was directly applied on the selected slices without any problem, despite the fact that both the scanners (GE vs. Siemens) and the resolutions (1.3mm $\times$ 1.3mm $\times$ 10mm vs. 1.7mm $\times$ 1.7mm $\times$ 6mm) were different. Hence, we obtain a set of 4D segmentations.

For each 4D image, one manual segmentation comes from the MICCAI challenge. This manual segmentation can be roughly described, for each 2D slice, as a smooth convex 2D curve that does not always follow the most contrasted structures seen in the images (see *e.g.* the green contours depicted in Fig. 2, first row). These manual segmentations are quite different from the ones drawn by our physicians on our own validation database (see *e.g.* Fig. 2, second row).

3.2 Evaluation method

In order to assess the methodology, the MICCAI Challenge proposes to compute the following measures : LV ejection fraction (EF), LV mass (LVM), and Dice index. EF and LVM are critical parameters for cardiac diagnosis and remodeling prevention. The EF is the amount of blood ejected during a heart cycle expressed

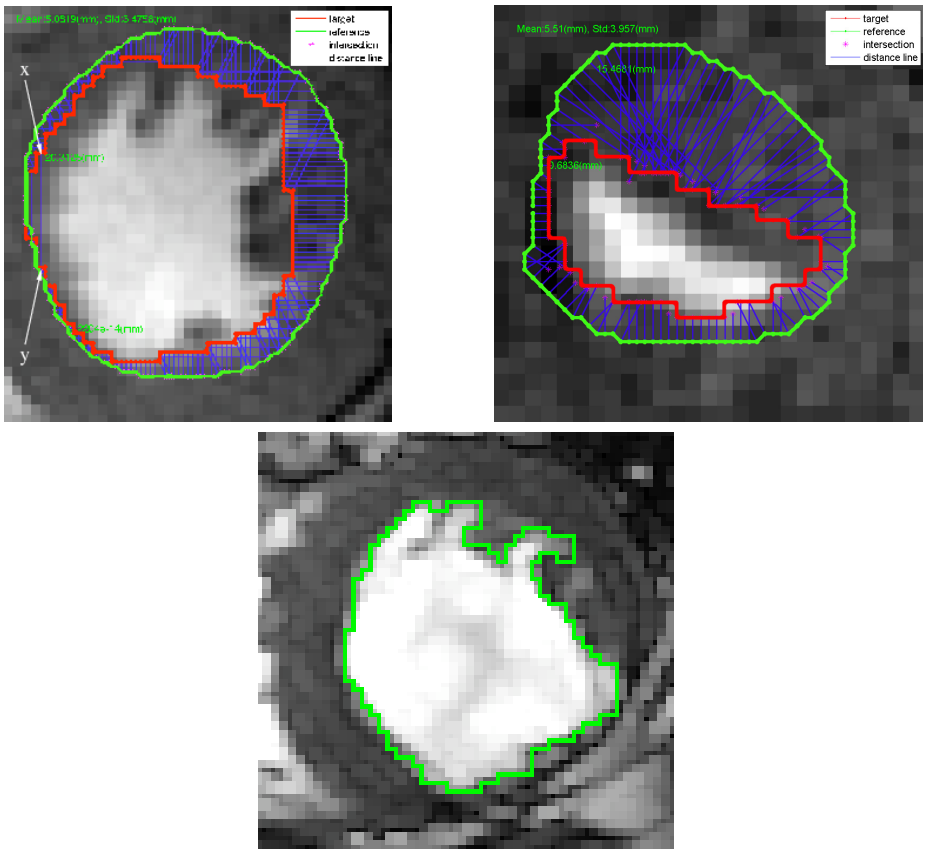

Fig. 2. First row: image examples from the challenge database. Second row: example of an image from our own database [2]. Green contours show the ground-truth segmentations of the endocardial borders. Red ones show the segmentations of the endocardial borders developed for this challenge (see Section 4.2).

as a fraction of the tele-diastolic volume. The Dice index is a measure of similarity between two sets. It ranges between 0 and 1: a value of 1 indicates a perfect agreement, whereas a value of 0 indicates an empty intersection. It is defined by $\frac{2\times|X\cap Y|}{|X|+|Y|}$, where X and Y are two sets (binary masks) representing two segmentations of a same object, and where $|X|$ denotes the size (are, volume, etc.) of set X.

Furthermore, the organizers of the challenge wanted to compute the average perpendicular distance (APD), that is meant to measure the distance from the automatically segmented contour to the corresponding manually drawn expert contour, averaged over all contour points. However, as it is implemented right now, the evaluation software used by the challenge organizers imposes that the

segmentation of each slice be made of a single smooth and closed curve. Thus the APD is not adapted to 3D surfaces, nor to discrete curves for which normals are difficult to compute. Indeed, the intersection of a 3D surface with a plane is not necessarily a 2D curve: it can consist of several closed curves and/or pieces of planes. Our method produces discrete (hence, non-smooth) 3D triangulated surfaces. This leads to great difficulties for evaluating our segmentation with APDs. For instance, in the left image of Fig. 2, it can be observed that the normals are not correctly computed (see in particular the normal emanating from point x). Because of such problems, the organizers decided to compute the normals emanating from the ground-truth instead of those emanating from our automated contours. Since the APD is not symmetric, this raises the problem of the comparison between our APD results and the APD results of the other participants, for whom the initial procedure was kept. We finally want to mention that the APD does not really reflect the distance between two curves. This can be seen for instance in the left image of Fig. 2, where for some points $\{x_i\}$ belonging to the green curve, the normals emanating from the individual x_i do not intersect the red curve at the point that is closest to that x_i.

Following this discussion, we decided to submit two sets of results for the challenge. The first one is the direct result of our algorithm, *i.e.*, binary masks of LV cavity and LV myocardium. The second one is a downgraded version of the first set, completed with missing slices, following the process described in the next Section 4. Only this second set has been considered by the evaluation team: indeed, due to the reasons we just discussed, the evaluation team was not able to quantitatively assess the first set, and only a qualitative evaluation was not deemed acceptable.

4 Adapting our method to the segmentation challenge

4.1 4D segmentation

We list below the different steps (and the associated user interactions) that our software requires to produce a 4D segmentation from a DICOM dataset.

1. Loading of the whole DICOM dataset in the software.
2. Manual selection of the slices that fulfill the guideline of Section 2.1.
3. Interactive adjustment of the image contrast (selection of two parameters).
4. Recording of a single mouse click in the center of the LV cavity on one 2D slice.
5. Registration of the adjacent 2D+t slices to form a single 3D+t sequence.
6. Segmentation of the 3D+t sequence.

Our segmentation method (*i.e.* step 6) comprises two kinds of parameters. A first series of four parameters is related to the geometry of the left ventricle. A second one, also made of four parameters, is related to brightness properties of cardiac MR images. The former has to be estimated only once. On the contrary, the later must be re-estimated for each new device since the brightness properties

of the left ventricle can change from one device to another. Thus to segment the images of the challenge, we only had to re-estimate the brightness-related parameters, which took less than one hour. Then, timings for performing the six steps described above range from 10 to 20mn.

4.2 From 4D surfaces to 2D curves

We downgraded the 4D segmentations obtained by our software in order to obtain curves that fulfill the evaluation format described in Section 3.2. To this end, we started by computing the intersection of each 2D slice and our 3D+t segmentation, leading, for each slice, to one binary 2D mask for the inside of the endocardial border and one binary 2D mask for the inside of the epicardial border. Then, we computed the convex hull of each mask from which we extracted the border edges. These sets of edges compose 2D curves that can be quantitatively assessed by the organizers of the challenge.

4.3 Missing slices

As shown above, the most basal and apical slices were discarded (see step 2, Section 4.1) before being processed by our software. In order to segment these, we: i) derive landmarks of the inside of the endocardial and epicardial borders from the already available segmentation of the adjacent slices; and ii) extract a minimal surface separating these landmarks from the borders of the 2D slices [9]. This constitutes the evaluated segmentation.

4.4 Quantitative assessment

We briefly summarize the results (sent by the organizers) of the quantitative assessment of the curves that we submitted. For the endocardial contours, our method achieved a mean shift of 3.0mm[1] $\pm$ 0.59 and a mean Dice index of 0.86 $\pm$ 0.04. For the epicardial contours our method achieved a mean shift of 2.6mm[1] $\pm$ 0.38 and a mean Dice index of 0.93 $\pm$ 0.01. The mean deviation of the ejection fractions (resp. left myocardium mass) computed from the automated segmentation with respect to the one obtained from the ground truth was 14% (resp. 23%).

We have not commented in details the challenge evaluation results. They look much inferior to those on our own dataset. We are unable at this time to assess the precise reason for this. We observe that on many slices our segmentation does not match the manual delineation. This may come from many factors, including perhaps medical informations we do not know. In order to obtain good segmentation results our experience is that a close collaboration with cardiologists is essential. This is of course very difficult to achieve in a challenge context. However, the Dice index achieved on the images of the challenge indicates that our method correctly segments these images.

[1] Note that all our contours were shifted of half a pixel (0.65mm) on both x and y axis, due to the use of different coordinate systems in our format and in the format of the challenge (see, *e.g.* the right image in Fig. 2).

5 Conclusion

Thanks to the challenge, we showed that our software runs on MR images acquired by GE scanners and also produces acceptable results in this situation. From our point of view, one of the most striking interest of such a challenge is to allow teams focusing on MR LV segmentation to meet and discuss together at the conference. We look forward to meetings all the people involved. We also wish to deeply thank the organisers of the challenge, in particular Dr. Perry Radau for his most extreme patience.

The challenge raises the question of the difficulty of having a general framework for evaluation, given the numbers of different methodologies set in different mathematical frameworks. The present challenge is a definite step forward in this direction. More generally, setting up such an evaluation framework for cardiac segmentation is precisely the goal of the French association IMPEIC. This association is a gathering of nine French teams working on cardiac segmentation. First results obtained by IMPEIC have been published in [10].

References

1. Cousty, J., Najman, L., Couprie, M., Clément-Guinaudeau, S., Goissen, T., Garot, J.: Automated, Accurate and Fast Segmentation of 4D Cardiac MR Images. In: Functional Imaging and Modeling of the Heart. Volume 4466 of LNCS., Springer-Verlag (2007) 474–483
2. Najman, L., Cousty, J., Couprie, M., Talbot, H., Clément-Guinaudeau, S., Goissen, T.: An open, clinically-validated database of 3D+t cine-MR images of the left ventricle with associated manual and automated segmentation. In: ISC/NA-MIC Workshop on Open Science at Medical Image Computing and Computer-Assisted Intervention. (2007) (9 pp.).
3. Beucher, S., Meyer, F.: The morphological approach to segmentation: the watershed transformation. E. Dougherty (Ed.), Mathematical Morphology in Image Processing, Marcel Decker (1993) 443–481
4. Vincent, L., Soille, P.: Watersheds in digital spaces: An efficient algorithm based on immersion simulations. IEEE PAMI **13**(6) (June 1991) 583–598
5. Cousty, J., Bertrand, G., Najman, L., Couprie, M.: Watershed cuts: Minimum spanning forests and the drop of water principle. IEEE PAMI **31**(8) (2009) 1362–1374
6. Cousty, J., Bertrand, G., Najman, L., Couprie, M.: Watershed cuts: Thinnings, shortest-path forests and topological watersheds. IEEE PAMI (2009) to appear.
7. Saito, T., Toriwaki, J.: New algorithms for Euclidean distance transformation of an n-dimensional digitized picture with applications. Pattern Recognition **27** (1994) 1551–1565
8. Bertrand, G.: Simple points, topological numbers and geodesic neighborhoods in cubic grids. Pattern Recognition Letters **15** (1994) 1003–1011
9. Appleton, B., Talbot, H.: Globally minimal surfaces by continuous maximal flows. IEEE PAMI **28**(1) (2006) 106–118
10. Kissi, A., Tilmant, C., De Césare, A., Comte, A., Najman, L., Lalande, A., Clarysse, P., Garreau, M., Sarry, L., Frouin, F.: Initiative multicentrique pour une plateforme d'evaluation en imagerie cardiaque. In: RITS 2009