RESEARCH ARTICLE

Comparative evaluation of active contour model extensions for automated cardiac MR image segmentation by regional error assessment

Duy Nguyen · Karen Masterson · Jean-Paul Vallée

Abstract *Objective* In the field of cardiac MR image segmentation, active contour models, or *snakes,* have been extensively used, owing to their promising results and to the numerous extensions proposed to improve their performance. This paper explores a methodology for evaluating cardiac MR image segmentation algorithms, which assesses the distance between computer-generated and the observer's hand-outlined boundaries. This metric was applied to various external force extensions of the traditional snake, since no systematic comparison has been performed.
Materials and methods Cardiac MRI from six patients were analyzed. Imaging was performed on a 1.5 T MR scanner with ECG-gated balanced steady-state free precession (b-SSFP) sequences. Segmentation performances were established for traditional snake, gradient vector flow snake, standard- and guided- pressure force-based snake. The use of a pre-treatment with non-linear anisotropic filtering was also compared to non-filtered images.
Results Agreement between manual and segmentation algorithms was satisfactory for ejection fraction for every segmentation scheme. However end-systolic and end-diastolic volumes were systematically underestimated.
Conclusion The developed regional error metric provided a more rigorous evaluation of the segmentation schemes in comparison to the classical derived parameters based on left ventricle volume estimation, usually used in functional cardiac MR studies. These derived parameters can furthermore mask local segmentation errors.

D. Nguyen (✉) · K. Masterson · J.-P. Vallée
Geneva University Hospital, Radiology Department,
1211 Geneva 14, Switzerland
e-mail: duy.nguyen@hcuge.ch

Introduction

Detailed information on cardiac dimensions and contractile function reflect the cardiac performance, whose quantification is of great interest because it can be used as a follow-up measure and as a prognosis indicator among patients suffering from heart diseases. Based upon this measurement depend significant diagnostic, prognostic, and therapeutic implications. A rapid, accurate, reproducible and non-invasive method of calculating the cardiac performance is desirable [1,2]. Cardiac magnetic resonance imaging (MRI) is a non-invasive technique and is considered as the best method to define cardiovascular anatomy. It has been established as the standard of reference for assessing ventricular function [3,4]. However, deriving physiological parameters from cardiac MR images requires extracting the endocardial and epicardial contours in a large number of diastolic and systolic images. This is a time-consuming task and a tedious manual procedure is required to obtain the quantitative results. This constitutes a limiting factor in the clinical use of cardiovascular MRI. In order to reduce the variability of the measurements and time constraints, a large number of automatic and semi-automatic procedures, based on various image processing approaches, have been proposed in the last few years. These procedures include thresholding and shape extraction [5], region growing [6,7], graph searching

[8,9], deformable models [10–13], loop B-spline curve fitting [14], and fuzzy clustering [15,16]. Among these methods, the active contour model, or *snake*, introduced by Kass [17] and representing a special case of the general multidimensional deformable model theory, has attracted most of the attention to date and has been extensively studied and used with promising results. By restricting the extracted boundaries to be smooth, the deformable models offer robustness to both image noise and boundary gaps. It also offers accommodation capacity to a significant variability of biological structures over time and across different individuals, as well as allowing the integration of the boundary elements into a coherent and consistent mathematical description. In order to improve the performance of this model, numerous extensions to the original snake formulation have been proposed. These extensions can be roughly classified in three main categories: geometrical representation, optimization algorithm, and external force extensions. Geometrical representation variations include B-spline snakes [18,19], T-snakes [20], and Fourier parametering [21]. Search algorithm variations mainly include dynamic programming [22], and the fast greedy-based algorithm [23]. The most well-known applicable external force extensions include the pressure force [24], the gradient-vector-flow (GVF) [25], and the distance potential force [26].

In spite of these various extensions, active contour models are still not implemented in a system intended for everyday clinical use. However clinical validation of these algorithms has been described in studies comparing automated and manual segmentations [12,13,27] with end-diastolic volume (EDV), end-systolic volume (ESV), and ejection fraction (EF) measurements. Graves et al. [13] showed that EDV and ESV were underestimated with semi-automatic methods, but there was no significant difference in the final calculation of EF. Van der Geest et al. [27] described smaller endocardial contours and larger epicardial contours with a semi-automated detection algorithm. Santarelli et al. [12] found a good agreement between manual and automatic left ventricle volume estimation with the use of a non-linear anisotropic filtering and a GVF snake. In spite of the promising results described, the relative merits of the different methods remain unclear and the optimal configuration of each method are in most cases unknown. One reason could be due to the lack of a gold standard for clinical data. Despite the intra- and inter-expert variability, manual drawing of the desired features by domain experts is actually the most used standard for comparison. Another reason is the difficulty in defining a metric which compares the computer-generated results and the segmentation results produced by expert observers. As described above, published validations of active contour model algorithms based on real cardiac MR images have used derived parameters from the extracted boundaries such as the areas enclosed by end-diastolic and end-systolic boundaries. The main drawback of these criterions is that local segmentation errors may be masked when using normalized or average parameters and therefore, comparing boundaries directly will provide a more stringent evaluation of the segmentation scheme [28,29]. Another reason justifying the need for more rigorous metrics is the progress made in the field of cardiac MRI in the last few years, allowing better spatial resolution [30,31] and therefore a better visibility of left ventricle papillary muscles and trabeculations. Hence, improvements in cardiac MRI mean more challenging segmentation difficulties for which an accurate method of segmentation is needed.

This paper addresses some of the problems identified above, firstly by exploring the relative merits of the principal snake extensions, considering that no attempt has been made to compare them in a systematic way. The methodology proposed differs from previous validation studies, in the use of a local metric applied on extracted features that would allow to assess segmentation errors in a more detailed manner than derived parameters from extracted boundaries. Direct assessment of the segmentation performances would allow an objective quantification of the computer-generated results by borders regions. This last issue is essential for example when quantification of the left ventricular wall thickness is desired for specific myocardial segments [32,33].

Materials and methods

Segmentation algorithms overview

This comparative study focuses on external force extensions of the active contour model and the filtering techniques usually used to improve cardiac MR images quality. The segmentation performances were established for a GVF snake (GVF), a pressure force-based snake (PF), and a guided pressure force-based snake (guided-PF) in comparison with the traditional snake (Trad) formulation. For the traditional model, the use of a pre-treatment by an anisotropic filter (Trad/Aniso) is also compared with the classical Gaussian blurring pre-treatment. The process of energy functional minimization of the extensions described above is based on dynamic programming because it has been reported as being optimal and numerically more stable and accurate [22]. In this section, the background of the various models that will be compared is briefly described.

Cardiac MR image filtering

In the classical schemes, the images first undergo some form of pre-processing operation. This is typically a low-pass filtering operation intended to reduce the effects of speckle noise. On MR images, noise is usually defined as a deviation from the true value considered representative of a tissue category. It has been shown that due to the non-linearity of the magnitude reconstruction process introduced to obtain real images, such simple assumptions about the image model are not realistic [34]. The transformation of complex MR data acquired in the Fourier domain to a magnitude image, changes the Gaussian noise distribution into a Rician distribution [35]. Hence applying conventional noise filtering schemes inevitably leads to biased results and can significantly reduce image contrast [36]. In such cases, the success in automated image segmentation of noisy images relies on the application of a suitable filtering technique that is able to remove noise in regions with homogeneous physical properties and to preserve or even enhance the edges. In the case of MR data, several authors have suggested the use of the non-linear anisotropic filter [12,37, 38] as proposed by Perona and Malik [37] in 1990. The Perona–Malik equation is a smoothing operation consisting of a diffusive process which suppresses boundaries by selecting proper spatial diffusion strength:

$$\frac{\partial I(x,y,t)}{\partial t} = \nabla c(x,y,t) \nabla I(x,y,t), \tag{1}$$

where ∇ is the gradient operator, $I(x,y,t)$ is the image intensity in function of spatial coordinates x,y and t is the iteration step. The term $c(x,y,t)$ is the conductance term which varies in space and time and is a monotically decreasing function of the magnitude of the gradient of the intensity:

$$c(x,y,t) = g(\|\nabla I(x,y,t)\|) = e^{-\frac{|\nabla I(x,y,t)|^2}{2K^2}}, \tag{2}$$

This function introduces a new parameter, K, which controls the influence of the gradient. In practice K acts as a threshold which determines whether to preserve edges or not: areas in which the ∇I is lower than K will be blurred more strongly than areas with a higher gradient magnitude. For this study, the standard anisotropic diffusion model has been implemented with the use of a single K value for both endocardial and epicardial borders, such as in the work of Santarelli et al. [12].

Traditional and guided pressure forces extensions

Cohen [24] has proposed increasing the attraction range by using a pressure force together with the Gaussian potential force. The advantage of using this additional force resides in its ability to search for contours from a rough, far away initialization. The pressure force is defined as:

$$F_p(v) = w_p N(v) \tag{3}$$

where $N(v)$ is the inward unit normal of the model at the point v and w_p is a constant weighting parameter. The main difficulty of the method lies in the selection of the weighting parameter value w_p. Its value has to be carefully chosen so that the pressure force is slightly smaller than the Gaussian potential force at significant edges, but large enough to pass through weak or spurious edges. In order to improve the traditional balloon model, an edge aggregating method was incorporated, allowing to guide the *pressure force* model. The aggregating method selected is based on the thresholding of image gradient by a sigmoid transfer function (STF). The STF is a smooth and continuous thresholding function of this type:

$$f(x) = \frac{1}{1 + e^{-ax}} \tag{4}$$

where for large a, the function's slope at $x = 0$ increases. To threshold the image gradient, (4) was adapted as follows:

$$f_{\nabla I}(x,y) = \frac{1}{1 + e^{\left(\frac{|\nabla I(x,y)| - c}{s}\right)}}, \tag{5}$$

where $|\nabla I(x,y)|$ is the gradient amplitude, c and s respectively define the STF center and slope. The center c establishes the threshold for which ∇I is considered significant. The slope parameter s establishes the threshold's contrast. As the absolute difference of ∇I values between homogeneous and boundaries regions will increase the absolute value of s will also increase. The choice of the STF centre c and slope s values depend mainly on image contrast. Optimal STF parameter values were determined according to image contrast, through the systematic assessment of threshold performances on cardiac-MR images by varying both STF centre and image dynamic range. The value of the slope parameter was fixed to 2 as it allowed a sufficient contrast between thresholded values. Threshold performances were assessed by calculating the difference between the average value in intra-tissues areas and in tissue boundaries through a profile line passing through both epicardial and endocardial borders. Then optimal STF centre was fixed at the maximal tissue value difference. The use of a positive slope parameter for the STF allows one to get a weighting parameter w_p (3) with maximal values in homogeneous regions and minimal values at the objects boundaries. The proposed guided pressure

force-based model is expected to be more robust, except in regions with little intensity gradient which might be further suppressed by the non-linear processing in Eqs. 4 and 5.

Gradient-vector-flow extension

Traditional snakes are known to be very sensitive to their initialization. Their second significant problem concerns difficulties in progressing into concave boundaries. The reason for this poor convergence is that the capture range of traditional potential force is generally narrowed toward the edge and more constant in homogeneous regions. Moreover, traditional potential force fields usually point horizontally in the opposite directions to the concave boundary portion. Thus, the curve is usually pulled apart and does not progress downward into the concave region. Although the pressure forces described above increase the capture range in an effective manner, they do not solve this second issue. Xu and Prince [25] proposed a new class of external forces that addresses this problem, for which fields are named gradient vector flow (GVF) fields. These are dense vector fields derived from images by minimizing an energy functional equation which diffuses the gradient vectors of a gray-level or binary edge map computed from the image. The GVF field is defined as the vector field $V(x, y)=[u(x, y), v(x, y)]$ that minimizes the energy functional:

$$E = \iint \mu \left(u_x^2 + u_y^2 + v_x^2 + v_y^2 \right) + |\nabla f|^2 |v - \nabla f|^2 dxdy, \tag{6}$$

where $f(x, y)$ is the edge map derived from the image $I(x, y)$ and ∇ is the gradient operator. This variational formulation is based on a standard principle with the objective of smoothing results when no data are available. Thus when $|\nabla f|$ is small, the energy is dominated by partial derivatives of the vector field, yielding a smooth field. On the other hand, when $|\nabla f|$ is large, the second term dominates the integrand, and is minimized by setting $v = |\nabla f|$. The parameter μ is a regularization parameter governing the trade-off between the first and second term, and usually set according to the amount of image noise. Once computed, the GVF field can be defined as a new static external force field $E_{ext} = v(x, y)$ and replaces the traditional potential energy function to obtain the corresponding dynamic snake equation.

Regional error metric

The best way to carry out a comparison between an automated segmentation and a group of experts' segmentations is so far unclear. Experimental evaluation of border identification algorithms is an aspect of this research which has been largely neglected. Various criteria have been used, including spatial overlap measurements, area or perimeter measurements which are usually used in cardiac MR image segmentation, and global boundary measurements such as the Hausdorff distance [39]. The main drawback of spatial overlap, area and perimeter measurements is the ability to show zero error between non-identical contours. The Hausdorff distance determines how far two finite points subsets of a metric space are from each other and does not take into account their geometrical nature. Therefore the Hausdorff distance between two curves which intersect each other many times will give a small distance, while the curves could be geometrically different. Metrics based upon regional comparison of the boundaries can provide a more precise evaluation of the segmentation scheme. The main difficulty of such metrics is the need for the establishment of artificial correspondence between the points on the two curves. One classical method to carry out this difficulty is the projection of radial lines from a common centroid [40,41]. Intersections of these radial lines with the two curves define the corresponding points, and the metric is defined as the root mean squared radial distance between all such points. However the weakness of this method is that all points of the boundaries need to be visible from the centroid and it fails for complex-shaped boundaries. In order to overcome this difficulty, we implemented a regional error metric based on a method proposed by Chalana and Kim [28], and which allows one to evaluate an average curve from two or more curves. The advantage of this method compared with other averaging methods is the establishment of one-to-one correspondence between the points of the curves to be averaged.

Given a set of N curves $C_1, \ldots, C_N$ each with P equidistant points, the first step of the method is to establish initial points of correspondence between the points x_{11} and x_{1P} on the curve C_1 and the respective closest points on the remaining curves, giving the sets of points $\{x_{21}, x_{2P} \{\ldots, \} x_{N1}, x_{NP}\}$. Points of the average curve C_a are given by the centroid of the N corresponding points:

$$x_{ai} = \frac{1}{N} \sum_{j=1}^{N} xji \quad \text{for} \quad i = 1, \ldots, P \tag{7}$$

The second step consists of computing the normal to every point of the average curve and finding its intersection with each of the N input curves. This new correspondence is averaged again to give a new average curve. The process is iterated until the average curve does not

Fig. 1 LV myocardium segmentation starting from the anterior junction of the RV with the LV, and propagation of the heart orientation through the stack of images covering the LV long axis
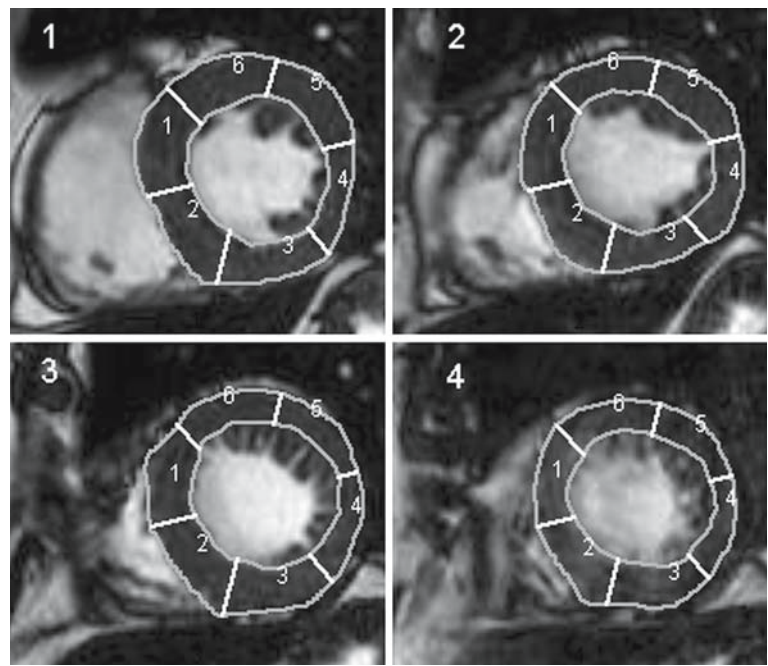
change anymore, which is usually done in less than five iterations. The use of the curve averaging method allows the evaluation of the regional difference between two given curves by computing the mean absolute Euclidian distance between the corresponding points. In order to compare the active contour model extensions described above, we measured the regional difference between the reference contour and the segmentation results according to each myocardial segment. To get a constant left ventricle (LV) myocardium division between all subjects, we applied a standardized myocardial segmentation defined by the American Heart Association (AHA) [42], which recommends six radial segments for the basal and mid-cavity short-axis slices and four radial segments at the apical level. These regions correspond to the areas of the myocardium supplied with blood from the different coronary arteries. According to the AHA recommendations, the anterior junction of the right ventricle (RV) with the LV is generally used as the reference to determine the heart orientation. For this study it was fixed manually on the first slice only of the image sequence covering the whole LV. The propagation of the heart orientation through the stack of images was done by converting the Cartesian coordinates of the manually fixed anterior junction of the RV with the LV into polar coordinates by the use of the reference contour centroid as origin. Then the heart orientation could be propagated slice by slice by computing the new RV anterior junction point according to the corresponding reference contour centroid, as shown in Fig. 1. This one illustrates the LV myocardium division through

the LV long axis. Myocardium segments are referred as follows:

- Segment 1 = Anteroseptal (AS) region
- Segment 2 = Inferoseptal (IS) region
- Segment 3 = Inferior (I) region
- Segment 4 = Inferolateral (IL) region
- Segment 5 = Anterolateral (AL) region
- Segment 6 = Anterior (A) region

In order to simplify the process, we kept six radial segments instead of four at the apical level. The AS and IS regions, as well as the IL and AL regions, being grouped together.

Data acquisition

Data from 11 patients referred for functional cardiac MRI in our center were analyzed. Data from five patients were used as separate training set in order to optimize the algorithm's parameters. Patients were aged between 25 and 87 years with a mean age of 46.1 years. There were no exclusion criteria except classical contraindications to MRI. Two patients presented a dilated cardiomyopathy, three patients had experienced a myocardial infarction prior to the MR examination, one patient presented a hypertensive cardiomyopathy, two patients presented a valvular cardiomyopathy, and three patients did not have any cardiac pathology. MR imaging was performed on a 1.5 T Intera MR system (Philips Medical System,

Best, The Netherlands). Long-axis reference views were used for positioning the necessary 5–8 perpendicular short-axis slices covering the entire LV. Images were collected during breath-hold with ECG-gated balanced steady-state free precession (b-SSFP) sequences and the following parameters: TR: 3.71 ms, TE: 1.86 ms, Flip angle: 70°, matrix: 160×256, FOV: 33 cm, slice thickness: 10 mm, NSA: 2, intersection gap: 10 mm. View-sharing reconstruction and parallel imaging SENSE [43] techniques were used to improve temporal resolution.

Image analysis

Parameters of the various segmentation models include the weighting constants of the internal energy term of the contour curve and can be defined as a measure of the model's elasticity and rigidity, respectively. The pressure force and GVF models include additional weighting parameters influencing the potential Gaussian force and the GVF forces, respectively. Parameters optimization was evaluated on a separate training set including five patients referred for functional cardiac MRI in our center. Once determined, all parameters were kept unchanged in the evaluation data set. For each subject, endocardial and epicardial contours were traced manually by one observer in end-diastolic and end-systolic phases. Papillary muscles and trabeculations were included in the LV cavity, according to the American Society of Echocardiography criteria [44]. All image segmentations were performed by a single trained observer with the use of a self-made application written in Java, without any manual correction done after semi-automated contour detection. The initial contour was placed only on the first slice of each phase as an oval region-of-interest (ROI) for each model. Image stacks were designed to start at the mitral valve level and to end at the apex level. The method of contour propagation between slices was chosen to be as simple as possible, in order to evaluate exclusively the capacity of each model to segment any desired feature changing in size and localization. Contours were propagated slice by slice as the segmentation result of the previous slice. On each segmented image, our regional error metric was applied. Classical derived parameters including EDV, ESV, EF and left ventricle mass were also calculated. Left ventricular volumes were measured as the sum of the cavity areas multiplied by the section interval (slice thickness + intersection gap). EDV was obtained from the first phase after the R wave. ESV was measured in the phase showing the smallest endocardial volume. EF was calculated as: ([EDV–ESV]/EDV) × 100%. The left ventricular mass was determined as the mean volume of the left ventricular wall multiplied by the specific

density of myocardium (1.05 g/cm^3) at both phases. In order to determine test-retest reproducibility, the above described classical derived parameters were measured at 6 months interval by a second trained observer on the same evaluation data set.

Statistical analysis

A statistical analysis of the measurements obtained with the implemented regional error metric and with the classical derived parameters EDV, ESV, EF and left ventricle mass has been performed. The agreement between results from manual and each of the segmentation models was expressed as mean difference $\pm$ SD of paired differences and was determined for EDV, ESV, EF and left ventricular mass. The percentage of difference was also calculated as the average difference between paired measurements divided by the mean of the two measurements. A two-tailed paired t-test was performed to determine the statistical significance of the observed differences. A P value less than 0.05 was considered to indicate statistical significance. The agreement between the segmentation models for the same derived parameters was also estimated, using the one-way analysis of variance (ANOVA). Bland-Altman [45] plots were used to analyze the differences between the different models. Reliability of the segmentation algorithms was estimated by the test-retest method with measurement of the change in the mean, typical error and intraclass correlation coefficient (ICC).

For the implemented regional error metric, the mean distance of each myocardial segment from the base to the apex of the LV between manual and semi-automated contours was averaged across all subjects. Results were reported as the mean distance $\pm$ SD according to each myocardial segment for both training and evaluation data sets. A two-tailed paired t-test was performed to determine the statistical significance of the observed differences between the training and evaluation data sets. In addition, one-way ANOVA was used to determine the significance of the difference between the various models according to each myocardial segment for the evaluation data set.

Results

The measurement of the optimal value for the guided-pressure force STF centre according to image contrast was estimated on eight different cardiac MR examinations. The results are displayed in the Fig. 2. Image contrast was determined as the maximal value of the image gradient module. The trend line of Fig. 2 shows
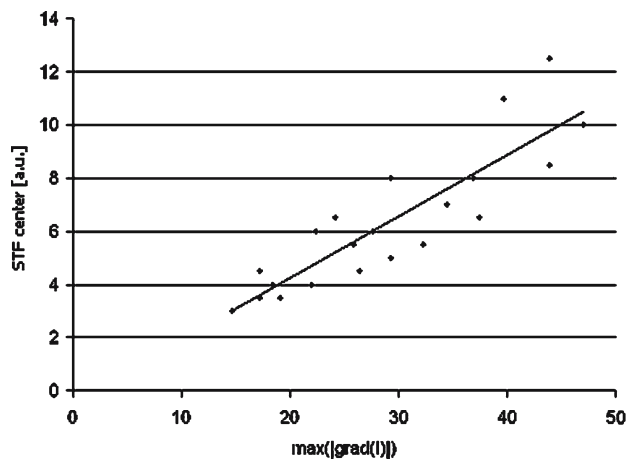
Fig. 2 Plot showing optimal STF center [(a.u) for arbitrary units] found at various image contrast, which is given by the maximal value of the image gradient module 141×105 mm (72 × 72 DPI)

that optimal STF centre value is linearly correlated to image contrast and can be evaluated by the following trend line equation:

$$c_{\text{optimal}} = 0.23 \cdot \max(|\nabla I|) - 0.39 \qquad (8)$$

Each of the segmentation models (Trad, Trad/Aniso, GVF, PF, guided-PF) was applied to both endocardial and epicardial contours in both end-diastolic and end-systolic phases, resulting in 480 segmented short-axis MR images. An example of segmentation difference is shown in Fig. 3 and the corresponding application of the local error metric is shown in Fig. 4. Mean Euclidian distances in pixels between the reference contour and segmented results for the training and evaluation data sets are shown in the plots of Figs. 5 and 6, respectively. The variations of mean Euclidian distances averaged on the six myocardial segments between the training and evaluation data sets for both endocardial and epicardial contours, are displayed in Table 1. These variations were not statistically significant ($P > 0.05$), except in the case of epicardium segmentation by the guided-pressure force model. Mean distance values in pixels for the evaluation data set are summarized in Table 2. The variations between the models tested were not statistically significant ($P > 0.05$), except for the anterolateral segment in end-diastolic phase ($P < 0.05$). For the endocardial border, mean Euclidian distances were in the range of 1.31–8.07 pixels, which gives a real distance ranging between 1.7 and 10.32 mm for a pixel resolution of 1.29 mm/pixel. For the epicardial border, there were no clear differences of performance between the various myocardial segments and the results were very poor in comparison with the endocardial border. The best segmentation results were performed with the traditional, the PF and the guided-PF models, with a statistically

significant difference in comparison with the segmentation schemes using the GVF model and a non-linear anisotropic filter. These results were obtained for almost every myocardial segment. Mean Euclidian distances were in the range of 2.22–10.08 pixels, which gives a real distance in the range of 2.9–13 mm for a pixel resolution of 1.29 mm/pixel. Finally, for both LV borders, the anterior myocardial region showed the worst results.

Mean differences and agreements between manual and automatic results (Manual-Automatic) for derived parameters measurement are summarized in Table 3. The table also shows inter-model agreement for the measurement of these parameters. The best agreements between manual and automatic measurements were obtained for the EF, except for the GVF model, which showed a statistically significant difference ($-7.3 \pm 5.71\%$; $P < 0.05$). Other derived parameters showed significant differences for every segmentation model, except for the traditional model with the use of a non-linear anisotropic filter. The guided-PF model showed the worst results in the estimation of ESV and EDV with a mean difference of $26.6 \pm 19.58\%$ and $25.21 \pm 11\%$, respectively. Concerning the LV mass estimation, the best results were also obtained with the use of the traditional model with anisotropic filtering ($5.85 \pm 16.45\%$; $P > 0.05$) and the worst results with the use of the pressure force-based model ($-28.17 \pm 18.83\%$; $P < 0.05$). Mean differences between manual and all segmentation models for the EF, LV mass, ESV and EDV measurements were -5.1, -6.31, 20.4 and 16.19%, respectively. Results from one-way ANOVA of inter-model variations did not show statistically significant differences of EF, ESV and EDV measurements ($P > 0.05$) while the LV mass estimation showed statistically significant differences between the methods ($P > 0.05$). In order to gain a more meaningful data interpretation, the systematic differences between manual and automated measurements are presented in the Bland–Altman plots in Fig. 7. It can be observed that ESV and EDV are systematically underestimated ($P < 0.05$) and this is true for all the segmentation models. On the contrary the LV mass tends to be overestimated ($P < 0.05$) by the traditional, pressure force-based and guided pressure force-based models.

Test-retest reliability measurements are displayed in Table 4. For every segmentation scheme, excellent reproducibility was obtained with EF and ESV measurements (ICC ≥ 0.95). EDV reproducibility was also satisfactory for every segmentation scheme (ICC between 0.8 and 0.97) but correlation coefficients were systematically lower compared to ESV measurements. Epicardial border segmentation also showed good reproducibility
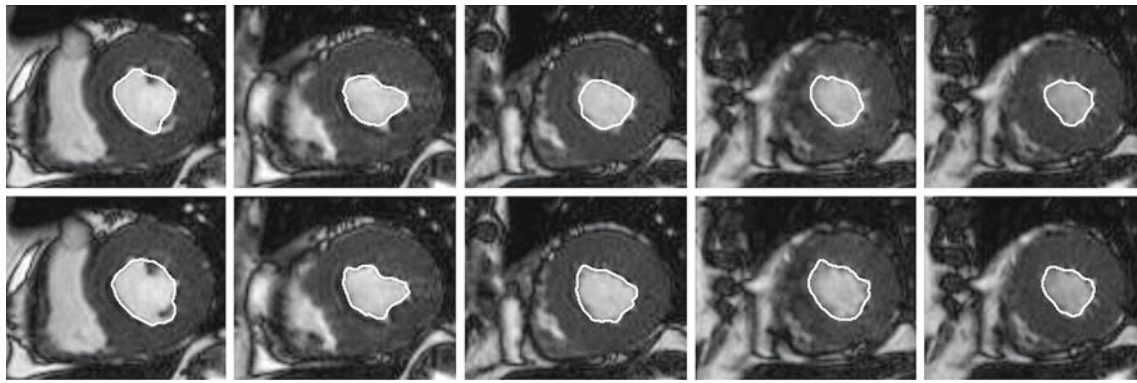
Fig. 3 Example of endocardial segmentation at the end systolic phase through the LV long-axis (*columns*) with the traditional model (*first row*) and the pressure forces-based model (*second row*)

Fig. 4 Example of regional error estimation between the manually traced contour (*grey curve*) and the contour found with the traditional model (*white curve*)

Fig. 5 Mean Euclidian distance in pixels of the training data set, according to LV myocardium segments and segmentation schemes. *Trad* Traditional snake, *Trad/Aniso* Traditional snake with anisotropic filtering, *GVF* GVF snake, *PF* Pressure force based snake, *Guided-PF* Guided Pressure force based snake 207×100 mm (72 × 72 DPI)
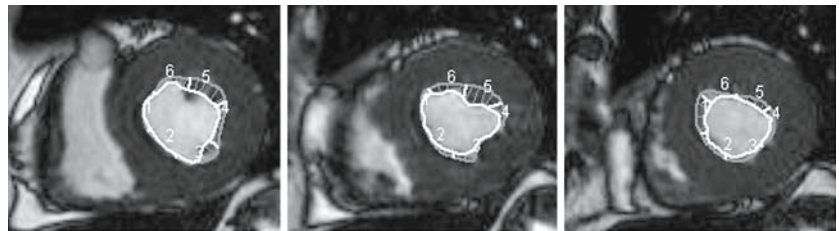

Fig. 6 Mean Euclidian distance in pixels of the evaluation data set, according to LV myocardium segments and segmentation schemes. *Trad* Traditional snake, *Trad/Aniso* Traditional snake with anisotropic filtering, *GVF* GVF snake, *PF* Pressure force based snake, *Guided-PF* Guided Pressure force based snake 207×100 mm (72 × 72 DPI)
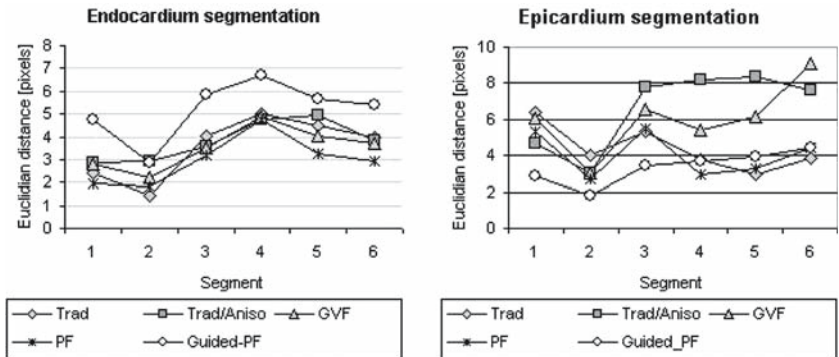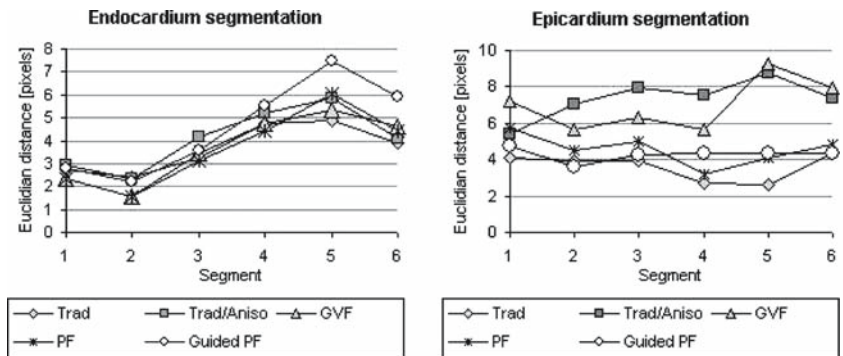
Table 1 Variation of mean Euclidian distances in pixels, averaged overall myocardial segments, between training and evaluation data sets, with corresponding P value (statistically significant if < 0.05)

	Endocardium		Epicardium	
	Change in mean $\pm$ SD [pixels]	P value	Change in mean $\pm$ SD [pixels]	P value
Trad model	0.1 ± 0.61	0.89	-0.8 ± 0.99	1.35
Trad/Aniso model	0.28 ± 0.52	0.68	0.72 ± 1.67	0.49
GVF model	0.14 ± 0.78	0.86	0.92 ± 1.67	0.37
PF model	0.67 ± 1.24	0.42	0.53 ± 0.76	0.41
Guided-PF model	-0.61 ± 1.55	0.55	0.93 ± 0.98	0.048

Trad Traditional snake, *Trad/Aniso* Traditional snake with anisotropic filtering, *GVF* GVF snake, *PF* Pressure force-based snake, *Guided-PF* Guided pressure force-based snake 185×48 mm (72 × 72 DPI)

Table 2 Mean absolute Euclidian distance $\pm$ SD in pixels, according to LV myocardial segment, segmentation model and LV border and phase

	Segment 1	Segment 2	Segment 3	Segment 4	Segment 5	Segment 6
Trad model						
Endocardium-Systole	3.07 ± 1.81	2.25 ± 0.86	2.96 ± 2.03	4.36 ± 1.92	5.4 ± 2.98	4.14 ± 2.54
Endocardium-Diastole	2.38 ± 1.08	2.6 ± 2.08	3.49 ± 1.47	5.16 ± 1.45	4.41 ± 1.17	3.62 ± 1.45
Epicardium-Systole	4.6 ± 1.72	5.08 ± 2.74	4.61 ± 1.52	3.19 ± 0.68	2.96 ± 0.54	5.45 ± 1.12
Epicardium-Diastole	3.59 ± 1.84	2.71 ± 1.15	3.22 ± 1.03	2.25 ± 0.9	2.22 ± 0.48	3.21 ± 0.48
Trad/Aniso model						
Endocardium-Systole	3 ± 1.76	2.17 ± 1.27	3.55 ± 2.21	4.71 ± 2.14	5.94 ± 1.82	3.81 ± 0.83
Endocardium-Diastole	2.82 ± 1.42	2.45 ± 2.65	4.74 ± 2.52	5.71 ± 3.67	5.79 ± 3.29	4.41 ± 2.43
Epicardium-Systole	5.12 ± 3.53	6.73 ± 2.64	8.26 ± 5.11	8.23 ± 6.32	9.02 ± 5.63	6.89 ± 3.13
Epicardium-Diastole	5.73 ± 3.61	7.05 ± 5.82	7.92 ± 5.75	7.57 ± 5.4	8.75 ± 6.03	7.35 ± 4.39
GVF model						
Endocardium-Systole	2.55 ± 0.57	1.5 ± 0.54	3.44 ± 1.74	5.45 ± 2.23	6.25 ± 3.85	5.24 ± 2.05
Endocardium-Diastole	2.1 ± 0.38	1.55 ± 0.37	3.4 ± 1.59	4 ± 0.9	4.38 ± 1.35	4.05 ± 1.91
Epicardium-Systole	7.97 ± 3.64	7.28 ± 4.31	6.27 ± 1.3	5.45 ± 2.28	10.08 ± 5.78	9.11 ± 2.72
Epicardium-Diastole	6.4 ± 4.18	3.98 ± 1.38	6.29 ± 1.76	5.81 ± 2.32	8.44 ± 5.66	6.72 ± 1.77
PF model						
Endocardium-Systole	2.65 ± 1.08	1.31 ± 0.54	2.44 ± 1.15	4.18 ± 2.41	6.37 ± 3.63	4.54 ± 2.77
Endocardium-Diastole	1.97 ± 0.99	1.84 ± 1.63	3.84 ± 1.71	4.7 ± 1.54	5.74 ± 2.26	4.29 ± 2.71
Epicardium-Systole	5.36 ± 2.81	4.57 ± 2.39	4.83 ± 1.64	3.24 ± 1.12	4.13 ± 2.19	4.3 ± 0.8
Epicardium-Diastole	6.13 ± 1.98	4.36 ± 2.02	5.17 ± 3.02	3.07 ± 1.59	4.12 ± 3.1	5.31 ± 2.63
Guided-PF model						
Endocardium-Systole	3.29 ± 1.13	2.51 ± 1.37	3.01 ± 1.28	4.89 ± 2.05	6.94 ± 3.67	5.81 ± 2.52
Endocardium-Diastole	2.32 ± 0.61	1.89 ± 0.47	4.14 ± 1.82	6.19 ± 1.71	8.07 ± 2.38	5.96 ± 1.94
Epicardium-Systole	5.05 ± 2.87	3.7 ± 1.88	3.76 ± 1.84	3.41 ± 1.48	3.95 ± 2.18	3.99 ± 2.37
Epicardium-Diastole	4.52 ± 1.6	3.5 ± 1.32	4.75 ± 2.9	5.24 ± 5.05	4.73 ± 3.89	4.76 ± 3.89
Inter-model SD						
Endocardium-Systole	0.31	0.52	0.44	0.5	0.57	0.81
Endocardium-Diastole	0.33	0.44	0.54	0.85	1.5*	0.88
Epicardium-Systole	1.34	1.5*	1.77	2.19*	3.27*	2.1*
Epicardium-Diastole	1.18	1.77*	1.64	1.94	2.77*	1.78*

Inter-model variations from one-way analysis of variance (ANOVA)
Trad Traditional snake; *Trad/Aniso* Traditional snake with anisotropic filtering; *GVF* GVF snake; *PF* Pressure force based snake; *Guided-PF* Guided pressure force based snake
*Difference statistically significant ($P < 0.05$)

for every segmentation model with a correlation coefficient for LV mass measurement varying between 0.77 and 0.96.

Discussion

The objective of the present study was to investigate two aspects of cardiac MR segmentation by active contour models that have been largely neglected. The first aspect concerns the relative merits of the various external force extensions proposed, since there has been no attempt to compare them in a systematic way. The second aspect concerns the method of evaluation of such algorithms, which is usually based on averaged parameters (EF, ESV, EDV and LV mass) and for which, local segmentation errors may be masked. This paper investigates a methodology allowing one to compare

Table 3 Mean difference between manual measurements and each of the segmentation model (manual-automatic) expressed in base units and as percentage

	EF (%)	LV Mass (ml)	ESV (ml)	EDV (ml)
Trad				
Base units	−0.69 ± 7.88	−48.33 ± 13.34*	20.79 ± 18.57*	28.42 ± 12.77*
Percentage	−1.81 ± 20.61	−20.59 ± 5.68	15.76 ± 14.08	14.23 ± 6.39
Trad/Aniso				
Base units	1.31 ± 10.37	11.97 ± 33.65	21.49 ± 23.58	34.24 ± 22.14*
Percentage	3.52 ± 27.85	5.85 ± 16.45	16.34 ± 17.92	17.4 ± 11.25
GVF				
Base units	−7.03 ± 5.71*	−36.02 ± 41.07	33.28 ± 26.3*	27 ± 15.61*
Percentage	−16.98 ± 13.8	−15.76 ± 17.97	26.49 ± 20.93	13.47 ± 7.79
PF				
Base units	−3.32 ± 8.31	−69.03 ± 46.15*	22.13 ± 17.15*	21.66 ± 19.42*
Percentage	−8.4 ± 21.02	−28.17 ± 18.83	16.87 ± 13.07	10.66 ± 9.56
Guided PDF				
Base units	−0.61 ± 6.97	−54.36 ± 49.07*	33.4 ± 24.59*	47.89 ± 20.89*
Percentage	−1.6 ± 18.26	−22.87 ± 20.64	26.6 ± 19.58	25.21 ± 11
Inter-model SD				
Percentage	8.07	12.39*	5.62	5.51

Inter-model variations from one-way analysis of variance (ANOVA) expressed as inter-model standard deviation
EF Ejection fraction; *ESY* end-systolic volume; *Trad* Traditional snake; *Trad/Aniso* Traditional snake with anisotropic filtering; *GVF* GVF snake; *PF* Pressure force based snake; *Guided-PF* Guided pressure force based snake
*Difference statistically significant ($P < 0.05$)

Fig. 7 Bland–Altman plots of the mean measurements obtained using the various segmentation schemes against their differences with manual measurements (Manual–Automatic). Plain lines show the mean and *dotted lines* show ±2 standard deviations. *EF* Ejection fraction; *ESV* end-systolic volume; *EDV* end-diatolic volume; *Trad* Traditional snake; *Trad/Aniso* Traditional snake with anisotropic filtering; *GVF* GVF snake; *PF* Pressure force based snake; *Guided-PF* Guided pressure force based snake 207 × 241 mm (72 × 72 DPI)
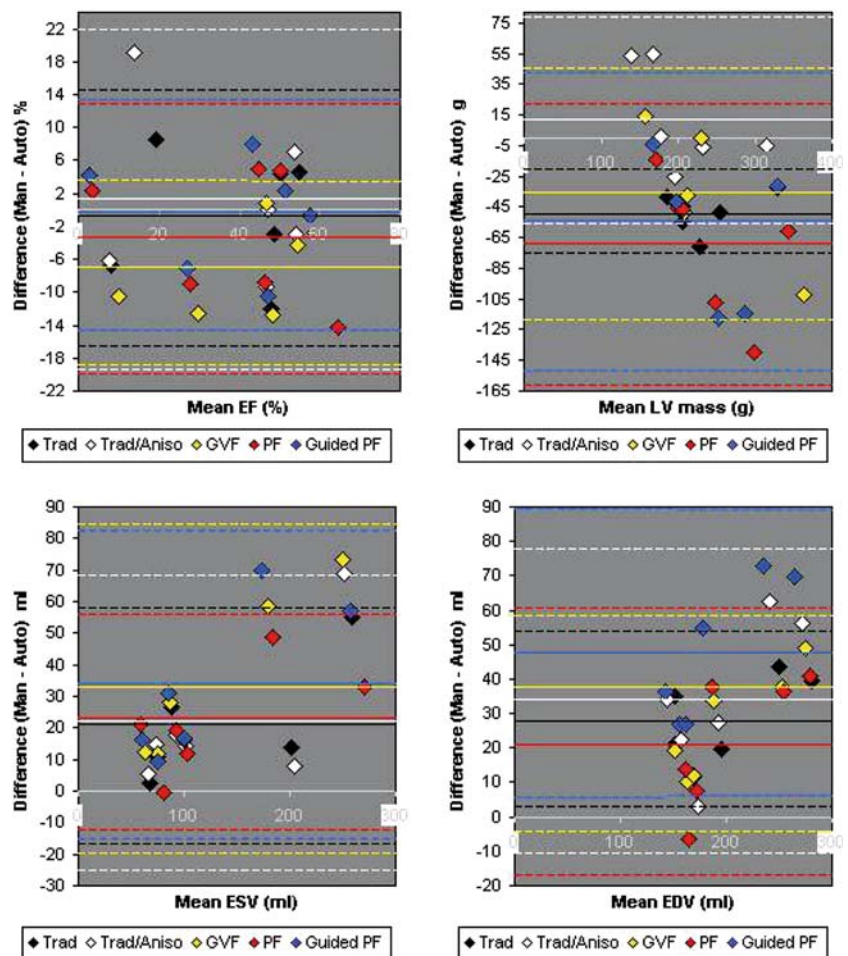
Table 4 Test- retest reliability

	EF %	LV MASS(ml)	ESV(ml)	EDV(ml)
Trad				
Change in Mean	5.09	31.59	−7.20	4.23
Typical error	4.98	20.45	5.82	12.82
ICC	0.95	0.93	0.99	0.94
Trad/Aniso				
Change in Mean	2.06	−25.74	−12.72	−18.11
Typical error	2.08	28.78	10.67	20.55
ICC	0.99	0.90	0.98	0.80
GVF				
Change in Mean	−2.97	15.14	−9.37	−26.06
Typical error	2.59	24.93	4.57	13.61
ICC	0.98	0.96	1.00	0.93
PF				
Change in Mean	4.29	13.61	−6.97	0.75
Typical error	5.44	20.50	6.72	8.11
ICC	0.95	0.93	0.99	0.97
Guided PF				
Change in Mean	1.39	61.24	−3.37	−3.43
Typical error	5.73	44.96	6.28	14.55
ICC	0.95	0.77	0.99	0.90

ICC Intraclass correlation coefficient, *EF* Ejection fraction; *ESV* end-systolic volume; *EDV* end-diatolic volume; *Trad* Traditional snake; *Trad/Ansio* Traditional snake with anisotropic filtering; *GVF* GVF snake; *PF* Pressure force based snake; *Guided PF* Guided pressure force based snake 159×141 mm (72 × 72 DPI)

boundaries directly and provides a more rigorous evaluation of the segmentation methods. The application of the regional error metric after LV endocardial segmentation process did not reveal any statistically significant difference in performance between the various external force extensions and filtering schemes. However for the epicardial border segmentation, statistically significant differences were found with poorer results for the GVF external force and also with the application of the traditional non-linear anisotropic filtering technique. Evaluation of the various segmentation schemes was also performed with traditional derived parameters. Good agreement between manual and semi-automated methods was found for the EF estimation, with the exclusion of the GVF model. ESV and EDV measurements were systematically underestimated with statistically significant differences with manually traced contours. On the contrary, the LV mass measurement was overestimated systematically except when pre-treatment with non-linear anisotropic filtering was used. The variations between the different segmentation methods were not statistically significant for EF, ESV and EDV measurements. Finally, every segmentation scheme was robust enough to show no statistically significant difference between the training and evaluation data sets, which had modified algorithm parameters. The reproducibility of the segmentation schemes was also investigated with the test-retest method with a 6 month interval between the two trials. Correlation coefficients were excellent for EF and ESV estimations, and satisfactory for EDV and LV mass estimations. Variations between the two trials could be mainly attributed to a different initial contour choice.

The use of the regional error metric allowed one to confirm objectively the key difficulties in cardiac MR segmentation. Theses difficulties include the complex endocardium borders, especially at the papillary muscle level and the lack of contrast between the epicardium and surrounding tissues. For the endocardial border, the metric clearly showed that segmentation results were less accurate in myocardial lateral and anterior regions due to the presence of papillary muscles and trabeculations respectively. These structures create a sufficient image contrast to prevent the snake from including papillary muscles into the LV volume estimation, and they cannot be blurred significantly with the use of traditional anisotropic diffusion technique. A method to carry out this challenge is the incorporation into the segmentation scheme of a priori knowledge [46–48] about the structure to be segmented. This allows one to fix a statistical shape into the global properties of the contour. However, while a priori knowledge integration significantly increases the segmentation robustness, one major limitation of these models is their inability to capture variability outside the learning set.

For the epicardial border, the segmentation performances of each model were less accurate and there was no clear difference between myocardial segments. This could be explained by the fact that each myocardial segment result was averaged on the distance found from the mitral valve to the apex level. Septal regions did not clearly show better accuracy because the contrast of the blood within the right ventricle usually disappears at the LV basal extremity. The regional error metric also showed that the GVF model and the use of a non-linear anisotropic filter gave the worst results for the epicardium, with statistically significant differences in comparison to others models. The GVF model is based on the diffusion of gradient vectors, in order to increase the capture range of image gradients. In the case of epicardial segmentation this can be a disadvantage, especially if the myocardium is thin (which is the case of patients suffering of dilated cardiomyopathy), because the model will be under the influence of both endocardial and epicardial borders, making it oscillate and resulting in poor segmentation performances. Concerning the non-linear anisotropic filter, poor epicardial performances could be partially explained by several reasons. The first one is the diffusion process proposed by Perona and Malik [38]. This process is recursive and it can introduce uncorrelated as well as unbounded gradients in case of image degradation by noise [49]. This makes it impossible to distinguish between edges which should be preserved and noise which should be diffused. Moreover, Catté et al. [49] have argued that the diffusion process is not well defined and it is not monotonic. In practice, this can lead to significantly different results for very similar images. The second reason may be the use of a single K value to filter both myocardial borders simultaneously, which is a disadvantage when the cardiac MR images are characterized by high variable gradient amplitude profiles at the endocardium and epicardium. Further works should consider more sophisticated filters such as the knowledge-based tensor anisotropic diffusion method proposed by Sanchez–Oritz et al. [50].

Correlations between the regional error and the use of derived parameters can be distinguished in two distinct cases. The first concerns EF, ESV and EDV which are derived from the endocardial border only. While a good agreement in EF measurements between manual and semi-automated methods has been found, ESV and EDV were underestimated with a statistically significant difference for each segmentation scheme. The underestimation of blood pool volumes can be explained by the exclusion of papillary muscles which can account for $6.5 \pm 1.3\%$ of the EDV [27]. This hypothesis has

been objectively demonstrated by the use of the regional error metric. The EF, being a ratio, this derived parameter will be less sensitive to systematic errors than absolute measurements, as long as they are consistent. The difference in estimation between the various segmentation schemes tested was not statistically significant for the EF, ESV and EDV measurements, which correlates with the results from the regional error metric. The second case concerns the LV mass estimation which is the only parameter based on both LV borders segmentation accuracy. The LV mass was found to be overestimated for every segmentation scheme, except with the traditional model after non-linear anisotropic filtering pretreatment. This overestimation suggests that epicardial contour were significantly larger with semi-automated methods and the worst results were found with the traditional, the pressure force-based and the guided pressure force-based models. A disagreement between the two metrics can be demonstrated: LV mass measurements showed best results with the use of the GVF model and the non-linear anisotropic filter, while these two models gave the worst results for the epicardial border segmentation with the regional error metric.

In comparison to others studies [12,13,27], the segmentation results were less accurate. However the tendency of semi-automated methods to underestimate endocardial volumes and to overestimate epicardial volumes, with no significant difference in the final calculation of EF, was confirmed. The marked differences between manual and semi-automated methods in the present paper can be explained by the absence of manual correction after automated segmentation [13,27], as well as no image exclusion due to automatic segmentation failure or unsatisfactory image quality [12]. The second reason is the method of propagation of the contour between slices. Our propagation method was as simple as possible in order to study exclusively the performances of the various models to segment any features changing in size and localization. However, because active contour models are very sensitive to the initial contour, it shows how much the method of contour propagation is important in a segmentation application based on active contour models. This aspect has largely been neglected and we believe that the segmentation success of an application in the field of cardiac MR images relies as much if not more on this aspect than on the use of deformable model extensions, for which the results demonstrated no significant differences in terms of performance. This point emphasizes the importance of an appropriate choice of slice thickness and intersection gap which should be as small as possible; in order to preserve myocardium borders continuity between slices.

Conclusion

In the field of cardiac MR images segmentation, active contour models have attracted most of the attention to date and have been extensively studied and used with promising results. However the exclusion of the papillary muscles and trabeculae from the myocardial wall still remains a segmentation challenge because of their MR image texture being very similar to the myocardium's texture. This paper differs from previously published validation studies on cardiac MR images segmentation, by the use of a regional error metric providing a more rigorous evaluation of the segmentation scheme. In comparison to the usual derived parameters based on LV volumes estimation, the proposed metric shows that such geometric assumptions can give wrong estimations on the real performances of the segmentation algorithm. While previous studies reported promising results through a good agreement between manual and semi-automated methods for the cardiac EF estimation, the local assessment of segmentation accuracy performed in the present study showed that there was still a number of segmentation difficulties at both endocardial and epicardial levels, despite the proposed extensions of the traditional snake formulation, and particularly despite the good agreement found with the EF estimation. Validation of segmentation algorithms based only on the use of normalized or average parameters may mask local segmentation errors and, therefore, methods of evaluation based directly on boundary discrepancy are required: firstly because they show better correlation with subjective assessment of the segmentation performances, and most importantly because objective and qualitative evaluation of various medical image segmentation algorithms is an important step toward their acceptance and clinical use.

References

1. Becker L, Silvermann K, Bulkey B, Kallman C, Mellits E, Weisfeldt M (1983) Comparison of early thallium-201 scintigraphy and gated blood pool imaging for predicting mortality in patients with acute myocardial infarction. Circulation 67:1272–1282
2. Pfeffer M, Braunwald E, Moye L, Basta L, Brown E, Cuddy T, Davis B, Geltman E, Goldman S, Flaker G, Klein M, Lamas G, Packer M, Rouleau J, Rutherford J, Wertheimer J, Hawkins C (1992) Effect of captopril on mortality and morbidity in patient with left ventricular dysfunction after myocardial infarction. N Engl J Med 327:669–677
3. Debatin J, Nadel S, Paolini J (1992) Cardiac ejection fraction: phantom study comparing cine MR imaging, radionucleide blood pool imaging and ventriculography. J Magn Reson Imaging 2(2):135–142
4. Higgins CB (1992) Which standard has the gold? J Am Coll Cardiol 19(7):1608–1609
5. Nachtomy E, Cooperstein R, Vaturi M, Bosak E, Veredb Z, Akselrod S (1998) Automatic assessment of cardiac function from short-axis MRI: procedure and clinical evaluation. Magn Reson Imaging 16(4):365–376
6. Singleton H, Pohost G (1997) Automatic cardiac MR image segmentation using edge detection by tissue classification in pixel neighborhoods. Magn Reson Med 37(3):418–424
7. Balzer P, Furber A, Cavaro-Menard C, Croue A, Tadei A, Geslin P, Jallet P, Le Jeune L (1998) Simultaneous and correlated detection of endocardial and epicardial borders on short-axis MR images for the measurement of left ventricular mass. Radiographics 18:1009–1018
8. Pope D, Parker D, Clayton P, Gustafson D (1985) Left ventricular border recognition using a dynamic search algorithm. Invest Radiol 26:295–303
9. Feagle S, Thedens D, Ehrhardt J, Scholz T, Skorton D (1991) Automated identification of left ventricular from spin-echo magnetic resonance images. Invest Radiol 26:295–303
10. Ranganath S (1995) Contour extraction from cardiac MRI using Snakes. IEEE Trans Med Im 14:328–338
11. Sebbahi A, Herment A (1997) Multimodality cardiovascular image segmentation using a deformable contour model. Comput Med Im Graph 21(2):79–89
12. Santarelli MF, Positano V, Michelassi C, Lombardi M, Landini L (2003) Automated cardiac MR image segmentation: theory and measurement evaluation. Med Eng Phys 25(2):149–159
13. Graves MJ, Berry E, Avedisijan A, Westhead M, Black RT, Beacock DJ, Kelly S, Niemi P (2000) A Multicenter Validation of an Active Contour-Based Left Ventricular Analysis Technique. J Magn Reson Imaging 12(2):232–239
14. Baldy C, Douek P, Croisille P, Magnin I, Revel D, Amiel M (1994) Automated myocardial edge detection from breathhold cine-MR images: evaluation of left ventricular volumes and mass. Magn Reson Imaging 12(4):589–598
15. Feng J, Lin W, Chen C (1991) Epicardial boundary detection using fuzzy reasoning. IEEE Trans Med Imaging 10(2):187–199
16. Lalande A, Jaulent M (1996) A fuzzy automaton to detect and quantify artery lesions from arteriograms. IPMU 1481–1487
17. Kass M, Witkin A, Terzopoulos D (1988) Active contour model. Int J Comput Vis 1:321–331
18. Menet S, Saint-Marc P, Medioni G (1991) B-Snakes: implementation and application to stereo. Artificial Intel Comput Vis 223–236
19. Brigger P, Hoeg J, Unser M (2000) B-Sline Snakes: a flexible tool for parametric contour detection. IEEE Trans Im Proc 9(9):1484–1496
20. McInerney T, Terzopoulos D (2000) T-snakes: topology adaptive snakes Med Image Anal 4(2):73–91
21. Staib L, Duncan J (1992) Boundary finding with parametrically deformable models. IEEE Trans Patt Anal Mach Intell 14(11):1061–1075
22. Amini A, Weymouth T, Jain R (1990) Using dynamic programming for solving variational problems in vision. IEEE Trans Pattern Anal Mach Intel 12(9):855–867
23. Williams D, Shah M (1992) A fast algorithm for active contours and curvature estimation. CVGIP: Image Understanding 55(1):14–26

24. Cohen L (1991) On active contour models and balloons. CVGIP: Imag Under 53(2):211–218
25. Xu C, Prince J (1998) Snakes, shapes and gradient vector flow. IEEE Trans Imag Proc 7(3):359–369
26. Cohen L, Cohen I (1993) Finite-element methods for active contour models and balloons for 2-D and 3-D images. IEEE Trans Pattern Anal Mach Intel 15(11):1131–1147
27. van der Geest R, Buller V, Jansen E, Lamb H, Baur L, van der Wall E, deRoos A, Reiber J (1997) Comparison between manual and semiautomated analysis of left ventricular volume parameters from short-axis MR images. J Comput Assist Tomogr 21(5):756–765
28. Chalana V, Kim Y (1997) A methodology for evaluation of boundary detection algorithms on medical images. IEEE Trans Med Imaging 16(5):642–652
29. Hammoude A (2001) An empirical parameter selection method for endocardial border identification algorithms. Comput Med Imaging Graph 25(1):33–45
30. Kuhl P, Spuentrup E, Wall A, Franke A, Schroder J, Heussen N, Hanrath P, Gunther R, Buecker A (2004) Assessment of myocardial function with interactive non-breath-hold real-time MR imaging: comparison with echocardiography and breath-hold cine MR imaging. Radiology 231(1):198:207
31. Moon J, Lorenz C, Francis J, Smith C, Pennell DJ (2002) Breath-hold FLASH and FISP cardiovascular MR Imaging: left ventricular volume differences and reproducibility. Radiology 223:789–797
32. Haag UJ, Hess OM, Maier SE, Jakob M, Liu K, Meier D, Jenni R, Boesiger P, Anliker M, Krayenbuehl HP (1991) Left ventricular wall thickness measurements by magnetic resonance: a validation study. Int J Card Imaging 7:31–41
33. Holman ER, Vliegen HW, van der Geest RJ, Reiber JH, van Dijkman PR, van der Laarse A, de Roos A, van der Wall EE (1995) Quantitative analysis of regional left ventricular function after myocardial infarction in the pig assessed with cine magnetic resonance imaging. Magn Reson Imaging 34:161–169
34. Gudbjartsson H, Patz S (1995) The rician distribution of noisy mri data. Magn Reson Med. 34(6):910–914
35. Sijbers J, Den Dekker A, Scheunders P, Van Dyck D (1998) Maximum likelihood estimation of Rice distribution parameters. Trans Med Imaging 17(3):357–361
36. Nowak R (1999) Based rician noise removal for magnetic resonance imaging. IEEE Trans Im Proc 8(10):1408–1419
37. Perona P, Malik J (1990) Scale space and edges detection using anisotropic diffusion. Trans Pattern Anal Mach Intell 12(7):629–639

38. Sijbers J, Den Dekker A, Van Der Linden A, Verhoye M, van Dyck D (1999) Adaptative anisotropic noise filtering for magnitude MR data. Magn Res Imaging 17(10):1533–1539
39. Huttenlocher D, Klanderman G, Rucklidge W (1993) Comparing images using the hausdorff distance. IEEE Trans Pattern Anal Mach Intell 15(9):850–863
40. Friedland N, Adam D (1989) Automatic ventricular cavity boundary detection from sequential ultrasound images using simulated annealing. IEEE Trans Med Imaging 10:344–353
41. Detmer P, Bashein G, Martin R (1990) Matched filter identification of left-ventricular endocardial borders in transoesophageal echocardiograms. IEEE Trans Med Imaging 9: 396–404
42. Cerqueira MD, Weissman NJ, Dilsizian V, Jacobs AK, Kaul S, Laskey WK, Pennel DJ, Rumberger JA, Ryan T, Verani MS (2002) Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart. Circulation 105:539–542
43. Pruessmann K, Weiger M, Scheidegger M, Boesiger P (1999) SENSE: sensitivity encoding for fast MRI. Magn Reson Med 42:952–962
44. Gottdiener J, Bednarz J, Devereux R, Gardin J, Klein A, Manning WJ, Morehead A, Kitzman D, Oh J, Quinones M, Schiller NB, Stein JH, Weissman NJ (2004) American society of echocardiography recommendations for use of echocardiography in clinical trials. J Am Soc Echocardiogr 17(10):1086–1119
45. Bland J, Altman D (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1: 307–310
46. Mitchell SC, Lelieveldt BPF, van der Geest RJ, Bosch HG, Reiber JHC, Sonka M (2001) Multistage hybrid active appearance model matching segmentation of left and right ventricles in cardiac MR images. IEEE Trans Med Imaging 20(5):415–423
47. Kaus MR, von Berg J, Weese J, Niessen W, Pekar V (2004) Automated segmentation of the left ventricle in cardiac MRI. Med Image Anal 8:245–254
48. Pluempitiwiriyawej C, Moura JMF, Wu YJL, Ho C (2005) STACS: New active contour scheme for cardiac MR image segmentation. IEEE Trans Med Imagin 24(5):593–603
49. Catté F, Lions P, Morel J, Coll T (1992) Image selective smoothing and edge detection by nonlinear diffusion. SIAM J Numer Anal 29:182–193
50. Sanchez-Ortiz G, Rueckert D, Burger P (1999) Knowledge-based tensor anisotropic diffusion of cardiac magnetic resonance images. Med Image Anal 3(1):77–101