Localization and Segmentation of Left Ventricle in Cardiac Cine-MR Images

Uday Kurkure, Amol Pednekar, Raja Muthupillai, Scott D. Flamm, and Ioannis A. Kakadiaris*

Abstract—Accurate delineation of the left ventricular myocardial boundaries on cardiac cine magnetic resonance (MR) images is essential for volumetric and functional cardiac analysis. Automated myocardial contour delineation often suffers from misalignment of slices, non-uniform coil sensitivity, blood flow related inter- and intra-slice intensity inhomogeneities, blurring due to motion, partial voluming and a need to circumscribe the papillary muscles and the trabeculae. In this paper, we propose a novel method for data-driven localization and segmentation of the left ventricle in the cine-MR images toward automated computation of ejection fraction. Our hybrid segmentation method combines intensity- and texture-based fuzzy affinity maps obtained from a novel multi-class, multi-feature fuzzy connectedness method with dynamic programming-based boundary detection to delineate the myocardial contours. Bland-Altman analysis indicates that the mean biases of the end-diastolic volume, end-systolic volume, and ejection fraction estimates of our method are comparable to the annotations from two experts.

Index Terms—Cardiac MRI, left ventricle, segmentation, fuzzy connectedness, dynamic programming

I. INTRODUCTION

Heart function is mainly determined by quantitative prognostic measurements of the left ventricle (LV) of the heart. This is achieved by delineation of the left ventricle myocardial boundary contours on a stack of two-dimensional (2D) short-axis (SA) cine magnetic resonance (MR) images covering the entire LV. Accurate estimation of the myocardial boundary is essential for reliable computation of critical LV functional descriptors (e.g., ejection fraction (EF)) in the diagnosis of various heart diseases. The manual contour tracing used in current clinical practice is labor-intensive, time consuming, and is subject to both intra/inter-operator variability and human error [1]. Research is ongoing toward the development of automatic cine-MR segmentation methods that could reduce the analysis time and more importantly have the potential to yield an accurate, unbiased and consistent method for myocardial contour delineation.

The detection of the myocardial boundaries requires several preliminary steps and each step has its own set of challenges. One of the first steps is to accurately identify and locate the LV in the SA images. This is an important step as it helps in providing good initialization for segmentation methods and also allows reducing the computational space by providing means to extract a region of interest (ROI) around the myocardium. However, this is a nontrivial problem and poses a number of challenges. First, due to image acquisition in multiple breathholds, the SA slices are prone to misalignment. Second, the orientation of the heart within the thoracic cavity is highly variable among different individuals, which prohibits simple assumptions about LV location (e.g., center of the image [2]).

The next step involves segmentation of the LV myocardial region, which poses the challenge of overcoming the in-plane blood signal intensity variation in the SA slices that can be attributed to heart motion and cardiac blood flow artifacts. Also, the volume stack of the SA slices contains through-plane intensity variations due to non-uniform coil sensitivity and is blurred due to partial voluming. Moreover, the delineation of boundaries of the anatomical structures composed of the same tissue type (e.g., myocardium, trabecular carneae and papillary muscles) and having similar $T_2/T_1$ values (e.g., for the liver and the myocardium in MR imaging) is challenging for intensity-based segmentation methods.

Several algorithms have been proposed to overcome the intensity- and shape-related challenges for a semi-automated and automatic segmentation of the LV. To be clinically viable, a segmentation method should be able to deal with pathologies (structural and functional variations) and routine imaging artifacts (intra- and inter-slice intensity variations). To achieve these objectives, we believe that a priori knowledge about the shape and image appearance of the LV is essential. However, the segmentation method should be driven by the individual characteristics of the image, rather than explicitly encoded constraints driven from a set of training samples. To overcome the challenges mentioned above, we propose a data-driven hybrid segmentation approach integrating intensity- and texture-based region segmentation with a dynamic programming-based boundary detection method to delineate myocardial contours.

In our previous work [3], we demonstrated methods for the localization of the LV using a temporal intensity difference map, the estimation of the LV blood pool region and the myocardial region using a greedy expectation maximization (EM) algorithm for Gaussian mixture learning [4] and segmen-
tation of the myocardium using intensity-based fuzzy affinity. However, our previous algorithm did not make use of scout-view images (which are routinely acquired to plan the short-axis views) for the LV localization, nor appearance information available in the short-axis images for tissue discrimination.

In this paper, we improve on our previous work and propose novel methods for the localization and segmentation of the LV towards automated computation of ejection fraction (parts of this work have appeared in [5]). Specifically, we propose a new method for the identification and localization of the LV in the SA slices using the intersection of the vertical long axis (VLA) and four chamber (4CH) scout views along with shape- and size-based geometric continuity constraints. We also present a novel formulation of the fuzzy connectivity region-growing segmentation method that integrates multiple intensity-based features with the topological relationship within and between the interrelated anatomical structures. The main attributes of this formulation are the following: 1) multi-class affinity that takes into consideration the affinity component distributions of all the neighboring regions (which allows competition between different objects); 2) multi-feature affinity that takes into account textural properties of the neighboring regions; and 3) Mahalanobis distance-based computation of affinities that integrates various affinity components without requiring the specification of weights.

The paper is organized as follows. In Section II, we provide a brief overview of the previous approaches toward cardiac segmentation in cine-MR. In Section III, we discuss in detail the steps of the proposed algorithm toward the segmentation of the myocardium and present our multi-class, multi-feature fuzzy connectedness formulation. In Section IV, we present the qualitative and quantitative results, and validate the method by computing geometric and clinical errors, followed by discussion in Section V and our conclusions in Section VI.

II. PREVIOUS APPROACHES

Previous approaches towards the myocardial boundary delineation for the volumetric and functional analysis of the cine-MR can be broadly classified into active deformable model-based approaches and hybrid segmentation approaches. Active model-based approaches aim to integrate image feature-derived information with a statistical shape or a probability distribution template. Hybrid segmentation approaches combine edge, region, and shape information for the extraction of the LV myocardial boundaries. The approaches to integrate intensity-based features and shape are based on, among others, fuzzy theory [6], dynamic programming [7], active contours (2D), deformable and variable shape (3D), appearance (3D) and motion (4D) models [6], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17] and hybrid segmentation approaches combining edge, region, and shape information [18], [2], [19]. The statistical approaches are ultimately limited by the sample size and the degree to which the training set is representative [15]. The 3D and 4D approaches are precluded for patients with inconsistent breath-hold due to the respiratory-induced slice displacements [11], [13], [20].

The common underlying principle behind the active model-based methods is the energy minimization between the image-derived features and a deformable model. While the initial active contour models constrained the energy minimization process based only on local 2D geometrical continuities [8], [9], recent models integrate statistical variations for landmark 3D global and local shape descriptors [6], coupled-shapes descriptors [11], global and local appearance features (e.g., intensity, gradient, texture) [10], spatially variant appearance features [11], motion (in-plane for cine-MR) descriptors [14], and coupled phases [15]. The variabilities of constraints are modeled either through significant modes of variation (PCA or ICA) or with non-parametric probability distributions and probabilistic atlases [12]. The main element of the shape and appearance modeling is building a good representative training data set and the requirement of correspondence across samples. The construction of the model from the training data is a tedious and laborious task because determining a point-to-point or tissue correspondence across subjects is challenging on a landmark-depleted myocardium. A deformable model, when it is initialized in the image volume, undergoes an iterative fitting process under the influence of image-derived features. The major challenge with the active models is to find an initial model, which is not too far from the desired image features, otherwise the model may deform to an incorrect local minimum. The initialization relies on some user interaction (e.g., selection of region of interest, seed point(s), initial contour(s) or explicit registration) or assumes the ventricles of near mean size to be located at relatively fixed location (e.g., center [2]) in the imaging plane. Simple grid-based initialization strategy requires determination of the required minimal grid size and spacing followed by multiple searches and then a way to determine which search has converged onto the true location of the heart [15]. In practice, the energy equation in most of the methods requires user-specified weight parameters for different energy components for an accurate model-to-image fitting [21]. The misalignment of the short-axis slices between breath-holds in cine-MR precludes the use of 3D or 4D models in patients with inconsistent breath-holds [11], [13], [20].

The MR data from different subjects differs in terms of anisotropic spatial resolution, number of slices and phases, and actual morphology-related positioning of the slices, rendering an essential step of data normalization across subjects for statistical models a difficult task. The interpolation along z axis (8-10 times the dimensions in x-y axes) and time (complex twisting and through-plane motion of the LV) or down-sampling in transverse dimension [15] may yield results different from the acquired data. The through-plane (due to non-uniform coil sensitivity) and in-plane (due to cardiac flow dynamics) intensity variation in cine-MR does not follow any specific pattern, thus further confounding the computation of intensity-based features required for an appearance model. The active shape, appearance, and motion models are ultimately limited by the sample size and the degree to which the training set is representative. Thus, the results could be biased toward a ‘too normal’ pattern of the LV shape and dynamics.

Research is ongoing in developing hybrid segmentation methods for the delineation of the myocardial contours by combining edge, region, and shape information [18], [2], [19].
Recently, in an approach based on varying tissue contrast in different pulse sequences, it has been shown that the LV segmentation based on 2D histogram clustering of dual contrast steady-state free-precession (SSFP) images yields quantitative information of global LV function comparable to that of manual delineation by experienced observers [22], [23]. The dual contrast approach seeks an optimal trade-off between the requirement of additional scans and making the LV analysis void of user interaction or parameter tuning. Considerable effort is still needed to optimize, compare, and validate each of these newer techniques.

Data-driven techniques need to be developed, which may work across the wide range of patient data. The anatomical objects in the imaging data are characterized by certain intensity level and intensity homogeneity features. The image elements seem to hang together to form a certain perception of the object region. Thus, medical image segmentation would benefit from a method based on the hanging togetherness property of the object of interest. The image segmentation framework based on fuzzy connectedness developed by Udupa and his collaborators [24] effectively captures the hanging togetherness of image elements specified by their strength of fuzzy connectedness. Connectedness, in general, consists of three major affinity components: 1) spel (spatial elements) adjacency, 2) intensity homogeneity, and 3) object-feature. One may devise a variety of functional forms for each component separately and combine them to the affinity relation suited for a specific application [25]. To effectively capture the degree of local hanging togetherness of the spels, both the homogeneity and object-feature-based components should be considered in the design of fuzzy spel affinities. However, in the previous approaches the homogeneity and object-feature based components were treated totally independent of each other. The functional forms and the parameters associated with these components allow many possible choices for fuzzy spel affinities. Saha et al. [25] conducted experimental studies for obtaining an insight into the best choices. The idea of using dynamic weights for affinity components was introduced by Pednekar et al. [26]. In our method, we modify the affinity formulation to incorporate multiple features to characterize complex appearances of the objects. We also introduce a sense of ‘competition’ among the neighboring objects to improve the segmentation by taking into account the appearance properties of the neighboring objects.

III. METHODS

This section provides a detailed description of our LV localization and segmentation methods. The proposed method comprises of the following steps: 1) localization of the LV and myocardium using scout-view geometry, blood-to-myocardial tissue contrast, and geometrical continuity constraints; 2) computation of myocardial region fuzzy affinity map using a multi-class, multi-feature fuzzy connectedness formulation to overcome the low contrast between the myocardium and the neighboring anatomical structures; 3) detection of myocardial boundary contours as optimal paths using dynamic programming to overcome anatomy-specific challenges, for instance same tissue types of the myocardium, the papillary muscles, the trabeculae carnea projecting out of the myocardium, and low liver-to-myocardium tissue contrast.

Before using our framework for the LV segmentation, we experimentally determine the relevant tissue types that are neighboring the LV myocardium. We also determine the relevant intensity-based textural features that can discriminate between the myocardium and the neighboring tissue types. This step is performed only once for this application. After this step, we obtain the relevant tissue classes and the intensity-based features that are used in our multi-class, multi-feature fuzzy connectedness formulation to generate a myocardial region fuzzy affinity map.

A. Data Acquisition

The studies were performed on five healthy volunteers (three males and two females), with a mean age of 30 years (range = 21-43 years), and fifteen clinical patients (twelve males and three females), with a mean age of 56 years (range = 38-70 years), all of whom were evaluated for left ventricular dysfunction. The study protocol was approved by the local Institutional Review Boards, and all subjects gave written informed consent. All subjects were imaged on a 1.5T commercial scanner (Philips Gyroscan NT-Intera) using a five-element phased-array surface "cardiac" coil and vector-cardiographic gating (VCG). The imaging protocol is described next. First, a multi-stack scout scan was obtained using a non-ECG gated steady state free precession (SSFP) technique. Using these single phase scout scans, a series of VCG gated cine SSFP images (TR/TE/flip: 3.2 msec/1.6 msec/60 deg; temporal resolution: 40 msec; acquired spatial resolution: 1.25 × 1.25 × 3 mm3) were acquired in the following order: a two chamber vertical long axis (VLA), a 4-chamber (4CH) view, and a series of contiguous short-axis slices covering the entire LV from the apex to the base (the level of the mitral valve annulus). Each cine slice was acquired during suspended respiration (6-8 secs/slice depending on the heart rate). All images were stored in the standard DICOM format.

B. Determination of relevant classes and features

In case of the SSFP cardiac MR data, the contrast between the blood and the myocardium is high. However, the tissue contrast between the myocardium and other neighboring anatomical structures is not adequate for intensity-based discrimination. Thus, before applying the LV segmentation framework, we experimentally determine relevant tissue classes neighboring the myocardium using intensity-based feature discrimination. We observed textural properties of the myocardium (muscles), LV blood, lungs (air), and liver. This step is required only once for a specific application. The relevant classes and features determined are then used to generate a fuzzy connected affinity map of a region of interest.

We computed Laws [27] and Gabor [28] texture features of the neighboring anatomical structures, namely myocardium (muscles), LV blood, lungs (air), and liver, to increase the discrimination. Laws texture features are generated by using
the 2D convolution kernels obtained by convolving five one-dimensional (1D) kernels with one another. Gabor features are obtained by filtering the original images with filter banks of four scales and six orientations of quasi-orthogonal Gabor filters. We used three mid-ventricular end-diastolic (ED) slices per subject from the MR data of ten subjects. In these 30 images, we manually delineated the myocardium, blood, liver, and air. It is important to note that we do not construct any shape/appearance model nor any feature space to be used during the segmentation; instead, our method constructs a feature space for each image slice individually during the segmentation process. The aim of this step is only to investigate the relevant features and classes for the region-based segmentation. To perform such an investigation, we need sufficient samples from each region. We used only the mid-ventricular slices in this step because the myocardial region is quite large in area in these slices, and thus, sufficient samples can be obtained from these slices for the feature analysis.

From the feature images we extracted pixel features for the following classes: the myocardium, the blood, the liver, and the lungs. The features were ranked according to their individual ability to discriminate myocardium from all the neighboring tissues using the Fisher’s criterion and the Mahalanobis distance measure. We found that spot-spot (ss) and spot-average intensity level (sl) Laws features are individually the most discriminating. Laws spot and level 1D kernels represent a second derivative filter and an averaging filter, respectively. The convolution of spot with spot, and spot with level kernels provide the ss and sl filters. Figure 1 depicts the Laws feature images for a short-axis image slice. Further, we found that combining intensity and intensity gradient of the ss feature provided higher discrimination between the myocardium and other neighboring tissues. However, the discrimination between the myocardium and the liver tissue remains low because the $T_2/T_1$ values of these tissues are very similar. Thus, the final feature vector (Eq. (2)) as used in our segmentation method consists of the pixel pair intensities, directional gradients, and the Laws ss feature.

C. Localization of LV

In cardiac cine-MR imaging, the short-axis view (Fig. 1(a)) is planned from the vertical long-axis (VLA) and the four chamber (4CH) scouts. This multi-view information, along with intensity, shape, and temporal information is used to detect the 3D medial axis of the LV in different temporal phases. First, we determine the centroid of the basal-adjacent slice in the end-diastolic (ED) phase by performing the following steps,

1) Compute the projection lines of the VLA and the 4CH scout images from ED phase to obtain an intersection cross-hair onto the end-diastolic short-axis slices using the offsets (0020, 0032) and angulations (0020, 0037) of the VLA and 4CH images obtained from standard DICOM header information (Fig. 2(a)).
2) Extract a ROI of the LV around the intersection cross-hair in the basal-adjacent slice. The ROI is required only to limit the search space. Thus, the closest approximation to the LV size is desirable but not required. From our observation of LV sizes in different subjects, we use approximately twice the largest LV size for the ROI.
3) Compute an 8-bit histogram of the ROI and threshold the ROI by applying the Otsu’s algorithm. The 8-bit normalization scales the images to the same range and also averages the spurious signals, thus providing a reliable Otsu’s threshold. The Otsu’s threshold is computed by maximizing a function of the ratio of the between-class variance and the total variance of the intensity levels. The assumption of a bimodal histogram by Otsu’s algorithm provides sufficient separation between the high intensity regions (blood) and the low intensity regions (muscles, liver, air).
4) Locate the LV in the threshold image by identifying the binary component that is closest to the cross-hair and is largest in area in the neighborhood as a rough approximate of the LV blood. Compute the centroid of the detected LV blood (Fig. 2(b)).

After locating the LV in the basal-adjacent slice of ED phase, the localization of the LV is then propagated in all subsequent short-axis slices of the same phase. However, the intersection cross-hair may not be always located within the LV due to misregistration of views, misalignment of SA slices, or curvature of the LV near the apex. To overcome this issue, the threshold blood region’s proximity to the intersection cross-hair, conformity to the estimated size and centroid location, and eccentricity are used as geometrical continuity constraints in a cost function for subsequent slices. The size and centroid of the LV in the current slice serves as the estimated LV size and centroid for the consecutive slices in the apical direction and provides the initial data for the progressive piecewise linear estimation of the size and centroid location of the LV.
the same as for the ED phase.

Intersection cross-hair, and the rest of the steps are repeated in the remaining slices. For other temporal phases, the ED LV polar image according to their radius value and clustering. The myocardial cluster is determined by finding the centroid whose value, and the last centroid, farthest from the origin and having the highest intensity clusters corresponding to the different structures. The first subtractive clustering formed by appending the rows of the polar image. We perform the feature $p, i$ and intensity signal $(i)$ (Fig. 4). The feature $p$ is computed as the position of a pixel in a vector formed by appending the rows of the polar image. We perform subtractive clustering [29] in the $(p, i)$ space to identify the clusters corresponding to the different structures. The first centroid, closest to the origin and having the highest intensity value, and the last centroid, farthest from the origin and having the lowest intensity value, are identified as the centroids for the LV blood and the air clusters, respectively. The centroid for the myocardial cluster is determined by finding the centroid whose intensity value forms a ratio of less than 0.5 with the intensity value of the centroid of the LV blood cluster. The threshold of 0.5 was determined experimentally and it holds because the cardiac cine-MR provides a high blood-to-myocardium intensity contrast. The clusters’ points are mapped back to the polar image to determine the seed pixels for the myocardium and the sample regions for the different structures. The sample regions are used to compute the feature distributions of different structures/classes for segmentation. Figure 5(a) depicts the cluster centers computed in the $(p, i)$ space but overlaid on the polar image. Similarly, Fig. 5(b) depicts the polar image with the overlaid sample region determined from the myocardial cluster in the $(p, i)$ space. Though the liver signals are present in the ROI, they are not included in the myocardial cluster. This can be attributed to the unique feature space of $(p, i)$. The attribute $p$ determines the closeness and the angular position of a pixel with respect to the origin. Since the myocardial signals are closer to the LV centroid and form a denser cluster than the liver signals, the clustering algorithm is able to obtain the desired myocardial cluster. Additionally, we select only the nearest neighbors of the myocardial cluster centroid in the $(p, i)$ space to further restrict the inclusion of the liver signals in the myocardial sample region. We found experimentally that the textural properties of air in the MR images do not change considerably across slices and subjects. Thus, we assume that it is safe to pre-compute the feature distribution of air and use it for all the subjects.

D. Determination of myocardium seeds and sample regions

Next, we determine the seed point(s) for the LV myocardium and the sample regions for all the classes. In the cine-MR data, the tissue types (due to their feature responses) and the LV myocardium (due to its spatial adjacency in polar coordinates) form clusters in a feature space, thus providing clues for LV blood and myocardium classification. We transform the short-axis images from the Euclidean coordinate system to a polar coordinate system (Fig. 3) and then construct a feature space of radial-angular position $(p)$ and intensity signal $(i)$ (Fig. 4). The feature $p$ is computed as the position of a pixel in a vector formed by appending the rows of the polar image. We perform subtractive clustering [29] in the $(p, i)$ space to identify the clusters corresponding to the different structures. The first centroid, closest to the origin and having the highest intensity value, and the last centroid, farthest from the origin and having the lowest intensity value, are identified as the centroids for the LV blood and the air clusters, respectively. The centroid for the myocardial cluster is determined by finding the centroid whose intensity value forms a ratio of less than 0.5 with the intensity value of the centroid of the LV blood cluster. The threshold of 0.5 was determined experimentally and it holds because the cardiac cine-MR provides a high blood-to-myocardium intensity contrast. The clusters’ points are mapped back to the polar image to determine the seed pixels for the myocardium and the sample regions for the different structures. The sample regions are used to compute the feature distributions of different structures/classes for segmentation. Figure 5(a) depicts the cluster centers computed in the $(p, i)$ space but overlaid on the polar image. Similarly, Fig. 5(b) depicts the polar image with the overlaid sample region determined from the myocardial cluster in the $(p, i)$ space. Though the liver signals are present in the ROI, they are not included in the myocardial cluster. This can be attributed to the unique feature space of $(p, i)$. The attribute $p$ determines the closeness and the angular position of a pixel with respect to the origin. Since the myocardial signals are closer to the LV centroid and form a denser cluster than the liver signals, the clustering algorithm is able to obtain the desired myocardial cluster. Additionally, we select only the nearest neighbors of the myocardial cluster centroid in the $(p, i)$ space to further restrict the inclusion of the liver signals in the myocardial sample region. We found experimentally that the textural properties of air in the MR images do not change considerably across slices and subjects. Thus, we assume that it is safe to pre-compute the feature distribution of air and use it for all the subjects.

E. Construction of myocardial affinity map

Having determined the seeds for the myocardium, and the feature distributions for the various structures in an SA image, next, we construct an affinity map of the LV myocardium using our multi-class, multi-feature formulation of fuzzy connectedness method. We describe our formulation using the terminology introduced by Udupa and Samarasekera [24]. An image is considered as a two-dimensional Euclidean space $R^2$ subdivided into spatial elements (spels or pixels in 2D). A digital space $Z^2$, where coordinates of a pixel correspond to a point, is set of all pixels of $R^2$. For any relationship $\phi$, the strength of $\phi$ between the two pixels $c$ and $d$ is represented by a membership function $\mu_{\phi}(c, d)$. If a fuzzy relation, $\alpha = \left\{ \left( (c, d), \mu_{\alpha}(c, d) \right) \mid c, d \in Z^2 \right\}$, is reflexive and symmetric, it is said to be a fuzzy spel adjacency, which describes the spatial location relationship between the two pixels. For any two pixels $c, d \in Z^2$, $\mu_{\alpha}(c, d)$ is assumed to be a hard adjacency relation, such that, $\mu_{\alpha}(c, d) = \begin{cases} 1, & \text{if } ||c - d|| \leq 1 \\ 0, & \text{otherwise} \end{cases}$, where $||c - d||$ represents the Euclidean distance between $c$ and $d$. The pair $(Z^2, \alpha)$, where $\alpha$ is a fuzzy spel adjacency is called a fuzzy digital space.

Local Fuzzy Spel Affinity: We define the local fuzzy spel affinity ($\mu_{\phi}$) to consist of three components:

a) the object feature intensity component ($\mu_{\phi}$),
b) the intensity homogeneity component ($\mu_\psi$), and
c) the texture feature component ($\mu_\varphi$).

This can be expressed in the following form,

$$\mu_\kappa(c, d) = \mu_\alpha(c, d)g(\mu_\psi(c, d), \mu_\varphi(c, d), \mu_\varphi(c, d)), \quad (1)$$

where $\mu_\kappa(c, d)$ is a hard adjacency relation. Thus, the fuzzy relation $\kappa$ in a digital space $Z^2$ indicates the degree of local hanging togetherness of pixels $c$ and $d$ in the space of the feature vector $x$, where

$$x(c, d) = \left[ \frac{1}{2} (f(c) + f(d)), \frac{1}{2} (f(c) - f(d)), \frac{1}{2} (t(c) + t(d)) \right]^\top. \quad (2)$$

Here $f(c)$ and $f(d)$ are the image intensities, and $t(c)$ and $t(d)$ are the texture features at pixels $c$ and $d$ in our framework.

The degree of membership of the feature vectors to an object of interest is determined by computing the distance of the feature vector from the feature distribution of the object of interest using the Mahalanobis distance metric $(D)$,

$$D^2(x(c, d)) = (x(c, d) - M_o)^\top \Sigma_o^{-1} (x(c, d) - M_o), \quad (3)$$

where $x(c, d)$, $M_o$, and $\Sigma_o$ are the feature vector, the mean feature vector, and the feature covariance matrix of the object of interest, respectively. The mean feature vector and the covariance matrix for different classes are computed from the sample regions determined from the detected clusters of different structures. The bias in intensity in a specific direction is accounted for by allowing different levels and signs of intensity homogeneities in different directions of adjacency [30]. Thus, this formulation accounts for different levels of the increase or decrease in intensity values in the horizontal or vertical directions. The advantage of using the Mahalanobis metric is that it weighs the differences in various feature dimensions by the range of variability in the direction of the feature dimension. These distances are computed in units of standard deviation from the mean. This allows us to assign a statistical probability to the measurement.

Global Object Affinity: Fuzzy connectedness captures the global hanging-togetherness of pixels by using the local affinity relation and by considering all possible paths between the two pixels (not necessarily nearby) in the image. It considers strengths of all possible paths between the given two pixels, where the strength of a particular path is the weakest affinity between the successive pairs of pixels along the path. Thus, the strongest connectedness path between the given two pixels specifies the degree of global hanging togetherness between them. Thus, the global object affinity is the largest of the weakest affinities between the successive pairs of pixels along the path $p_{cd}$ of all the possible paths $P_{cd}$ from $c$ to $d$ and is given by

$$\mu_K(c, d) = \max_{p_{cd} \in P_{cd}} \left\{ \min_{1 \leq i \leq m} \left[ \mu_\kappa(c(i), c(i+1)) \right] \right\}.$$ 

Global Class Affinity: In our framework the global object affinity and local pixel affinity are assigned only if the membership value of $c$ and $d$ belonging to the neighboring objects’ classes is less than a pre-defined value (note that the membership value has inverse relationship with Mahalanobis distance metric in our formulation). The neighboring objects are defined as the objects with common boundaries in Euclidean space. For a given pixel pair $(c, d)$, we compute a discrepancy measure with respect to the predetermined or existing distributions (covariance matrices) of the neighboring classes in terms of its Mahalanobis distance. Then, the minimum discrepancy measure, $J(c, d) = \min_{1 \leq i \leq b} D(x(c, d))$, where $b$ is the number of neighboring classes of the target object, provides the maximum membership value of a pixel pair belonging to a certain class. If the $J(c, d) < \varepsilon$, and the class to which the pixel pair belongs is not the target object class, then the local pixel affinity $\mu_{\kappa(c,d)}$ is set to zero, else the pixel pair is a candidate to be considered to belong to the target object, hence its local pixel and global object affinity is computed as described earlier. We determined the membership threshold $\varepsilon$ experimentally to be 3 for this application.

**Single Seed vs. Multiple Seeds:** One significant advantage of the fuzzy connectedness formulation as proposed by Udupa and Samarasekera [24] is that the algorithm can be initiated from multiple seeds without any computational penalty. This can be attributed to the dynamic programming approach used in this formulation allowing computation of optimal paths in a non-exhaustive manner initiated from multiple starting points. We take advantage of this property and use multiple seeds (here, four) for the construction of the fuzzy connectedness map of the myocardium. The seeds are selected from the myocardial cluster that was obtained in the previous step, as mentioned in Section III-D, such that they are located at angles of $\pi/4$, $3\pi/4$, $-3\pi/4$, and $-\pi/4$ from the LV centroid in the Cartesian coordinate system. The use of multiple seeds overcomes the thinning of the myocardium and the variations in the signal intensity values within the myocardium. The obtained fuzzy map provides soft segmentation of the myocardium assigning values between 0 and 1 to the pixels. A hard segmentation could be obtained from the fuzzy map by thresholding it using Otsu’s method. Figure 6(a) depicts such a fuzzy map obtained by the multi-class, multi-feature fuzzy connectedness method using a single seed for the polar image. Figure 6(b) depicts the binary segmentation of the fuzzy map obtained by the Otsu’s method. It can be observed that the region-based segmentation is not able to completely separate the papillary muscles, liver, RV myocardium, and trabeculae (in apical slices) from the myocardial region because of the similar signal properties.
The gradients of the myocardium affinity image provide boundaries located almost in the horizontal direction in the polar coordinates. The optimal myocardial contour is the contour which follows the high myocardial affinity and high myocardial affinity gradient closely, while maintaining a high degree of spatial continuity in the tangent direction. Such continuity constraints make it feasible to separate the papillary muscles, liver and other structures from the myocardial region. The dynamic programming approach is used to detect the myocardial boundary by determining a horizontal optimal path between the two ends of the polar image.

Any possible boundary can be represented as a polyline with $N$ vertices $(P_1, P_2, P_3, \ldots, P_N) \in P$. For a polyline to be a valid boundary, it should have minimum value for the cost function, $C_{sum} = \sum_{j=1}^{N} C(P_j)$. The cost function for the polar myocardial affinity image is expressed as

$$C(P_j) = -\omega_i C_i(P_j) - \omega_g C_g(P_j) - \omega_r C_r(P_{j-1}, P_j),$$

where $\omega_i$, $\omega_g$, and $\omega_r$ are the weights for the myocardial affinity value $C_i$, the myocardial affinity gradient $C_g$, and the radial distance between pixels on the polyline in adjacent columns $C_r$, respectively. The affinity value term forces the boundary to follow a homogeneous path through the pixels with high myocardial affinity. The affinity gradient term is responsible for moving the boundary toward the points having strong myocardial affinity gradient in a direction perpendicular to the boundary. The continuity term restricts the boundary from taking big steps in the radial direction between the pixels in the adjacent vertical columns. Thus, it imposes the spatial continuity constraint, smoothing out the boundary in the horizontal direction. This continuity term can force a first- or a second-order continuity constraint. The second order constraint allows for smoother transitions of the boundary. The boundary detection by dynamic programming is computationally efficient because it obviates the need for an exhaustive search of the optimal solution. The weights for the cost components of the cost function are obtained experimentally.

Figure 7(a) depicts the endocardial boundary contour obtained by the dynamic programming method in a polar transformed image. One advantage of our method is that by reversing the direction of the gradient kernel, we can also determine the epicardial boundary of the myocardium (Fig. 7(b)). The detected boundary contours are then transformed back to the original coordinate system to obtain circular myocardial boundary contours.

**F. Compute myocardial boundary contours**

The data were transferred to a separate post-processing workstation (EasyVision, Rel. 5.0, Philips Medical Systems). An experienced CMR expert (14 years in CMR) and another clinician from St. Luke’s Episcopal Hospital reviewed the images and manually drew contours on the short-axis SSFP cine images of ED and ES phases. The manual annotations were performed on the same ED and ES phases. For consistency, the following definitions were used throughout the manual analysis process. ED was defined as the phase revealing the largest cavity area and ES was defined as the phase with the smallest cavity area. The basal slice was defined as the most basal slice where the ED LV blood volume was surrounded by 50% or more of ventricular myocardium. In order to study the effects of anatomical structure on the accuracy of the algorithm, each slice was also annotated as basal (from the mitral valve annulus to the tips of the papillary muscles at the ED), mid-cavity (the entire length of the papillary muscles), or apical (beyond the papillary muscles to the apex). The LV outflow tract (LVOT) was also included in the basal slices for accurate volumetric and ejection fraction computation purposes.

The proposed method is implemented in MATLAB, and takes about four minutes to process one phase volume on a Windows machine with a dual core, 2.21 GHz AMD Athlon processor. The endocardial and epicardial contours determined by our algorithm for Subject-1 are depicted in Fig. 8. These qualitative results demonstrate that our method can reliably detect the myocardial boundary contours for both ED and ES phases. However, the goal of this paper is to obtain the endocardial boundary accurately for the EF computation and thus, we present the quantitative validation for the LV localization and the endocardial boundary detection only.

To assess the accuracy of the LV localization method, the centroid ($c_m$) and the mean radius ($r_m$) of the manually annotated boundary contours are determined first. The centroid $c_m$ of the manual contour ($B_m$) is treated as the gold standard for the LV location for validation purposes. Next, we compute the distance ($\delta$) between the centroids obtained manually ($c_m$) and by our method ($c_\alpha$) as explained in Section III-C. A
The Dice similarity coefficient (DSC) was used to quantify the overlap between the manual and the automatically generated contours. The DSC is defined as

$$DSC(S_m, S_a) = \frac{2|S_m \cap S_a|}{|S_m|+|S_a|}$$

where $S_m$ and $S_a$ represent the manual and automatic contours, respectively, and $|X|$ denotes the number of pixels in set $X$. A DSC value of 1 indicates perfect overlap, while a value of 0 indicates no overlap.

To assess the segmentation quality, we computed the relative square distance error (RSDE). RSDE is defined as

$$RSDE = \frac{\delta^2}{r^2}$$

where $\delta$ is the distance between the estimated center of the LV and the manually annotated center, and $r$ is the radius of the manual contour. A smaller RSDE indicates better segmentation accuracy.

The Bland-Altman analysis was used to assess inter-observer variability. The mean biases and variability for the LV volumes and ejection fraction are listed in Table I. The results show good agreement between the automated and manual segmentation methods, with mean biases of less than 2% for most volumes and a regression value of 0.899 for the ejection fraction.

In conclusion, our method achieved good overlap and accuracy in delineating the LV contours in 2D image slices from ED and ES phases. The method produced accurate results for the LV volumes and ejection fraction, with mean biases of less than 2% and a high degree of correlation between manual and automatic annotations.

Table I: Bland-Altman Analysis Results for the LV Volumes and EF.

<table>
<thead>
<tr>
<th>Volume</th>
<th>$AR_1$ Bias</th>
<th>$AR_2$ Bias</th>
<th>$R_1R_2$ Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>ED-Basal (ml)</td>
<td>-3.11</td>
<td>0.19</td>
<td>3.29</td>
</tr>
<tr>
<td>ED-Mid (ml)</td>
<td>-7.89</td>
<td>0.19</td>
<td>4.35</td>
</tr>
<tr>
<td>ED-Apical (ml)</td>
<td>-5.03</td>
<td>-3.34</td>
<td>3.82</td>
</tr>
<tr>
<td>ES-Basal (ml)</td>
<td>-1.60</td>
<td>-0.60</td>
<td>1.00</td>
</tr>
<tr>
<td>ES-Mid (ml)</td>
<td>-5.96</td>
<td>-3.37</td>
<td>2.58</td>
</tr>
<tr>
<td>ES-Apical (ml)</td>
<td>-1.26</td>
<td>0.17</td>
<td>1.43</td>
</tr>
<tr>
<td>EDV (ml)</td>
<td>-16.02</td>
<td>-2.96</td>
<td>7.57</td>
</tr>
<tr>
<td>ESV (ml)</td>
<td>-8.82</td>
<td>-3.80</td>
<td>6.99</td>
</tr>
<tr>
<td>EF (%)</td>
<td>1.57</td>
<td>1.83</td>
<td>5.70</td>
</tr>
</tbody>
</table>

The agreement in terms of DSC was $85.53 \pm 12.34\%$, indicating good overlap between our results and the manual annotations. The Bland-Altman analysis also confirmed the good accuracy of our method compared to the manual annotations. This study demonstrates the potential of our method for automatic segmentation of LV contours and accurate estimation of LV volumes and ejection fraction.
V. DISCUSSION

LV functional assessment using the short-axis cine-MR images involves a large amount of data (e.g., it is common to collect about 10-14 slices, each with 25 to 30 temporal points or phases during the cardiac cycle). While recent advances in acquisition methodology and MR imaging hardware make it possible to readily acquire such large amounts of data, the analysis of this data has largely remained manual. This is further complicated by the fact that cine-MR images may be blurred as a result of cardiac motion, patient motion, background signal variation in MR due to coil fall off, and intrinsic partial voluming effect. Our method effectively overcomes these challenges by combining domain-specific knowledge using VLA and 4CH scouts with a dynamic programming-based region growing algorithm which takes multiple classes and textures into consideration. The main advantage of our approach over other existing techniques is that it is completely data-driven and does not require creating any kind of shape or appearance model. The local appearance statistics are computed for each image slice individually. Thus, we avoid the problems of differences in the training and the unseen data. Another major advantage of our method is that it locates the LV without requiring any kind of manual initialization.

With the current research focused on 3D/4D segmentation methods of anatomical structures, we argue that our proposed 2D approach is still relevant for the mentioned application. The input to the most 3D methods is an MR volume constructed by stacking the 2D SA slices from different spatial location but same time point. Such a reconstructed 3D volume may not represent actual geometry if the slices are misaligned due to patient motion in same breath-hold and different inspiration levels in multiple breath-holds. Although methods [31] for movement correction of SA slices have been proposed, such a post-processing step incurs a time penalty because it requires additional slices in long axis views. Secondly, the blood signal intensity varies from base to apex with no predictable pattern [32] due to the variable coil intensity rendering the 3D segmentation difficult for the 3D methods. These challenges underline the need for a 2D slice-by-slice approach for the localization and the segmentation of the LV in the cine-MR images. Moreover, the 3D/4D segmentation approaches make certain global shape and appearance assumptions. Due to the global nature of these assumptions, the segmentation averages over the differences between slices at different position. This is one the main reason that the existing 3D/4D methods are not able to effectively separate the trabecular carneae and the attached papillary muscles from the myocardium. Using a 2D approach, we avoid these problems and obtain segmentation in close agreement with the manual annotations. However, our method could be extended to 3D/4D by combining the 2D segmentation component with 3D/4D global connectivity component. A global higher dimensional model could be fitted to the left ventricle by deforming the boundary surface using the detected myocardial boundary in 2D slices. Such combination of local 2D approach and global 3D/4D approach would allow the separation of the trabeculae carneae and the papillary muscles, and also improve the myocardial segmentation globally by taking advantage of both spatial and temporal connectivity.

VI. CONCLUSION

In this paper, we proposed a novel approach toward the computation of the ejection fraction from cardiac cine-MR data. Specifically, we introduced methods to locate the LV and determine its boundary in short-axis slices. We also presented a novel region growing segmentation method based on the fuzzy connectedness framework that incorporates information from different classes and textural features. We evaluated our method on routine clinical data sets of asymptomatic and symptomatic subjects. The results of our method were evaluated against manual annotations by the experts in terms of the errors in the boundary estimation, area estimation and ejection fraction estimation. The results of our method were within the limits of inter-observer variability and consistent with the manual tracings by the experts. The mean biases and the variability of our method for the LV volumes and the ejection fraction were comparable to the mean biases and the variability in the inter-observer variability analysis. Thus, the performance of our method indicates a high potential to be used in the clinical setting for computation of ejection fraction.

ACKNOWLEDGMENT

This work was supported in part by NSF IIS-0431144, IIS-9985482, IIS-0335578, and CNS-0521527. Any opinions, findings, conclusions or recommendations expressed in this material are the authors’ and may not reflect the views of the NSF.

REFERENCES


10


Uday Kurkure received the Ph.D. degree from the University of Houston, Houston, TX, in 2008. He is currently a postdoctoral fellow in the Computational Biomedicine Lab, Dept. of Computer Science at the University of Houston, Houston, TX. His research interests include medical image analysis, computer vision, and pattern recognition.

Amol Pednekar received the B.Engg degree from Victoria Jubilee Technical Institute, University of Mumbai, in 1996. He received the Ph.D. degree in computer science from University of Houston, Houston, TX, in 2003. From June 2004 to October 2005, he was a Radiology Research Specialist with Cardiovascular MRI research group at St. Lukes Episcopal Hospital and Texas Heart Institute. Since November 2006, he has been a MR Clinical Scientist with Philips Medical Systems, Cleveland, OH. His research interests are in cardiac magnetic resonance imaging and analysis. These include work on image processing, and cardiac segmentation and motion analysis.

Raja Muthupillai

Scott D. Flamm serves as Head of Cardiovascular Imaging within the Imaging Institute at the Cleveland Clinic, and holds joint appointments in Cardiovascular Medicine and Pediatrics. He is a member of multiple professional societies, serves or has served on the Editorial Boards of Circulation, Radiology, International Journal of Cardiovascular Imaging, Journal of Cardiovascular Magnetic Resonance, and the Texas Heart Institute Journal, and has authored or co-authored over 100 scientific publications. He has also been elected to the Board of Trustees of the Society for Cardiovascular Magnetic Resonance, and the North American Society of Cardiac Imaging. Dr. Flamm has developed a busy clinical service in cardiovascular CT & MRI, and continues to perform research devoted to improving patient care.

Ioannis A. Kakadiaris (M’91) received the B.Sc. degree in physics from the University of Athens, Athens, Greece, the M.Sc. degree in Computer Science from the Northeastern University, Boston, MA, and the Ph.D. degree from the University of Pennsylvania, Philadelphia. In August 1997, he joined the University of Houston (UH), Houston, TX, after a Postdoctoral Fellowship at the University of Pennsylvania. He is currently the Eckhard Pfeiffer Professor of Computer Science, Electrical & Computer Engineering, and Biomedical Engineering. Dr. Kakadiaris is the founder of the Computational Biomedicine Laboratory and he is also the 2008 Director of the Methodist-University of Houston-Well Cornell Medical College Institute for Biomedical Imaging Sciences (IBIS). His current research interests include cardiovascular informatics, biomedical image analysis, biometrics, computer vision, and pattern recognition. He is the co-inventor of in vivo vasa vasorum detection using differential imaging. Prof. Kakadiaris is the recipient of a number of awards, including the National Science Foundation (NSF) Early Career Development Award, the Schlumberger Technical Foundation Award, the UH Computer Science Research Excellence Award, the UH Enron Teaching Excellence Award, and the James Muller Vulnerable Plague Young Investigator Prize. His research has been featured on Discovery Channel, National Public Radio, KPRC NBC News, KTRH ABC News, and KHOU CBS News.

Authorized licensed use limited to: TUNG HAI UNIVERSITY. Downloaded on March 10, 2009 at 21:57 from IEEE Xplore. Restrictions apply.