

Cardiac anchoring in MRI through context modeling

Xiaoguang Lu^{1*}, Bogdan Georgescu¹, Marie-Pierre Jolly¹, Jens Guehring¹,
Alistair Young², Brett Cowan², Arne Littmann³, and Dorin Comaniciu¹

¹Siemens Corporate Research, Princeton, NJ, USA

²Auckland MRI Research Group, University of Auckland, Auckland, New Zealand

³Magnetic Resonance, Siemens Healthcare, Erlangen, Germany

Abstract. Cardiac magnetic resonance imaging (MRI) has advanced to become a powerful diagnostic tool in clinical practice. Robust and fast cardiac modeling is important for structural and functional analysis of the heart. Cardiac anchors provide strong cues to extract morphological and functional features for diagnosis and disease monitoring. We present a fully automatic method and system that is able to detect these cues. The proposed approach explores expert knowledge embedded in a large annotated database. Exemplar cues in our experiments include left ventricle (LV) base plane and LV apex from long-axis images, and right ventricle (RV) insertion points from short-axis images. We evaluate the proposed approach on 8304 long-axis images from 188 patients and 891 short-axis images from 338 patients that are acquired from different vendors. In addition, another evaluation is conducted on an independent 7140 images from 87 patient studies. Experimental results show promise of the proposed approach.

1 Introduction

In cardiology, precise information on both the dimensions and functions of the heart chambers is essential in clinical applications for diagnosis, prognostic, and therapeutic decisions. The precision on the measures extracted from MR images has been demonstrated and makes MR imagery a standard for left ventricle (LV) analysis [1]. Although cardiac MR imaging technologies have rapidly advanced [2, 3], due to considerable amount of available data, analysis of cardiac images for quantification is time consuming and error-prone for human operators.

Typical cardiac MR studies contain both long-axis and short-axis slices. Long axis slices are not only used as scout images for acquisition planning, but also are complementary to the short axis stack [4]. Long axis slices capture heart chamber shape information and can also be used to correct mis-registration of the short axis stack. Anchoring is helpful for accurate and efficient cardiac modeling, such as initialization of deformable model based approaches [5], accelerating acquisition time by facilitating fully automatic planning of cardiac MR examinations, and accurate assessment of mass and volume [6], where demarcation of

* Corresponding author (xiaoguang.lu@siemens.com).

the base is important [7]. For example, anchoring the base plane and the apex in long-axis images facilitates and accelerates the LV segmentation of the short stack. Anchoring RV insertion (intersection between RV outer boundary and the LV epicardium) helps analyze LV functions according to AHA myocardial segmentation models [8]. Each anchoring component can be of different geometric representations, e.g., a line segment for the base plane and a point for the apex.

We propose a unified approach to detecting anchoring components. Anchoring components are converted into parameterized bounding box representations, which fit into an object detection framework. Such representation embeds not only individual anchoring components but also their context, which contains rich information to distinguish the anchoring components from its background and other anatomical structures. We apply a learning-based method to train detectors on expert annotations in order to handle complex appearance and heterogeneous characteristics of anatomical features in medical images, as the complex prior knowledge is implicitly encoded. Learning based object detection approaches have been demonstrated successful in many applications [9, 10].

The proposed approach provides a large flexibility to be applied to a wide range of anatomical structures. We apply our framework to detect LV base mitral valve plane and LV apex in long-axis images, and RV insertions and RV lateral (a point where the RV outer boundary changes directions significantly within the image) in short-axis images, as shown in Fig. 1. Our approach is fully automated.

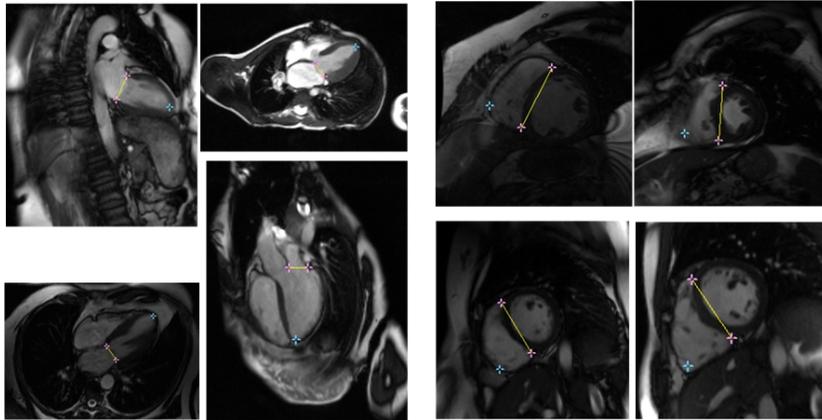


Fig. 1. Examples of cardiac images and associated anchoring components of interest.

2 Methodology

We propose a unified framework to anchor the anatomy of interest. We convert detection of different types of anchoring components, such as base plane (a line) and apex (a point) for LV in each long-axis image, into the same object detection framework by designing a contextual representation for each anchoring

component. Each anchoring component is represented by a bounding box as an object with 5 parameters (2 translations, 1 orientation, and 2 scales), as shown in Fig. 2. A probabilistic learning approach [11] is applied to solve a two-class classification task, i.e., object vs. background. In order to reduce computational cost of searching through a large 5-dimensional parameter space, we adopt the marginal space search strategy proposed in [12].

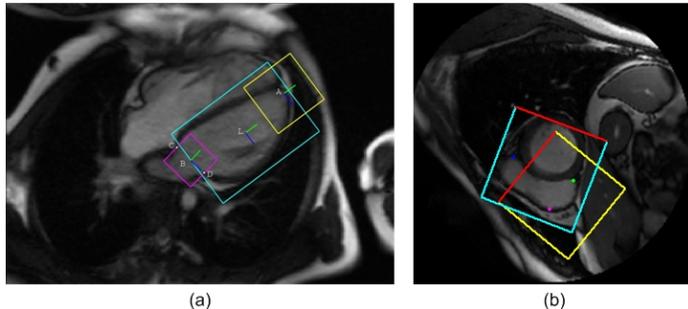


Fig. 2. Context construction and conversion from anchoring components to objects represented by a parameterized bounding box. Base plane (pink) and apex (yellow) are shown in a long-axis image (a). RV insertion (cyan) and RV lateral (yellow) are illustrated in a short-axis image (b). Notice that each bounding box is parameterized by its positions, orientation (green axis as local x-axis in (a), and red edge indicating orientation in (b)), and scales. By adjusting these parameters, a different context around the anchor component is selected.

2.1 Context learning

A 2D object (bounding box) parameter set consists of five degrees of freedom. Exhaustively searching in this 5-dimensional space is prohibitive for online applications. Therefore, we adopt the marginal space search strategy, where we design a series of detectors that estimate plane parameters at a number of sequential stages in the order of complexity, i.e., translation, orientation, and scale, as the parameter degrees of freedom increase [12]. Different stages utilize different image features. Multiple hypotheses are maintained between algorithm stages, which quickly removes false hypotheses at the earlier stages while propagating the right hypotheses to the final stage. Only one hypothesis is consolidated as the final detection result.

We use a probabilistic boosting tree [11] for each detector to achieve a discriminative task between the object and background. The classifier is a tree-based structure with which the posterior probabilities of the presence of the object of interest are calculated from given image data. Therefore, each detector not only provides a binary decision for a given sample, but also a confidence value associated with the decision. The nodes in the tree are constructed by a combination of simple classifiers using boosting techniques [11].

Each detector selects a set of discriminative features that are used to distinguish the object from background from a large pool of features. For the classifiers at the translation stage, we choose Haar wavelet-like features [9], which are efficiently calculated using integral image-based techniques. For the classifiers at the orientation and scale stages, steerable features [12] are applied, because their computation does not require image rotation and re-scaling, which are computationally expensive, especially when the hypothesis search space is large.

2.2 Context modeling in long-axis slices

For each long-axis image, two anchoring components are targets of interest, the base plane and the apex. We associate a two-dimensional bounding box with each contextual object around the anchoring component. Each bounding box is specified by a five-parameter set θ , containing two positions $\langle x, y \rangle$, one orientation $\langle \phi \rangle$, and two scales $\langle s_x, s_y \rangle$. Although only positions are mostly in use, orientation and scales are useful in encoding proper and consistent context learned during offline training process, where a set of contextual models/classifiers are obtained.

We collect a set of cardiac long-axis images and annotate the anchoring components to learn contextual models. Based on this annotated training set, we build a contextual model for each target object. For long-axis images, a joint contextual model [13] for the pair of $\langle \text{apex, base plane} \rangle$ is also constructed as shown by the cyan box in Fig. 2, which is used for inter-anchoring validation. Let $\langle x_a, y_a \rangle$, $\langle x_{b1}, y_{b1} \rangle$, and $\langle x_{b2}, y_{b2} \rangle$ denote the positions of the apex, and two basal annulus points, respectively. The contextual parameter set for the base plane is: positions $\{(x_{b1} + x_{b2})/2, (y_{b1} + y_{b2})/2\}$; orientation $\{\text{orthogonal to the line segment connecting the two basal annulus points, and pointing toward the apex side}\}$; and scales $\{s_b, s_b\}$, where $s_b = \sqrt{(y_{b2} - y_{b1})^2 + (x_{b2} - x_{b1})^2} * \alpha$, where α is a factor that can be used to adjust the contextual range set to 2.4 in our experiments. Selection of α is a tradeoff between rich context and noise. For the apex, the context parameters are constructed as: positions $\{x_a, y_a\}$; orientation $\{\arctan((y_a - (y_{b1} + y_{b2})/2)/(x_a - (x_{b1} + x_{b2})/2))\}$; and scales $\{s_a, s_a\}$, where $s_a = \sqrt{((y_{b2} - y_{b1})^2 + (x_{b2} - x_{b1})^2)} * \alpha$.

2.3 Context modeling in short-axis slices

For each short-axis image, two anchoring components are targets of interest, the RV insertion and the RV lateral. For RV insertion, the anterior and posterior anchors are identified through the following context modeling. Let $\langle x_{an}, y_{an} \rangle$, $\langle x_{po}, y_{po} \rangle$, and $\langle x_{La}, y_{La} \rangle$ denote the positions of the RV insertion anterior, RV insertion posterior, and RV lateral, respectively. The contextual parameter set for RV insertion is: positions $\{(x_{an} + x_{po})/2, (y_{an} + y_{po})/2\}$; orientation $\{\text{orthogonal to the line segment connecting RV insertion anterior and RV insertion posterior, and pointing toward the LV}\}$; and scales $\{s_{ins}, s_{ins}\}$, where $s_{ins} = \sqrt{(x_{an} - x_{po})^2 + (y_{an} - y_{po})^2} * \beta$. β is set to 1.5 in our experiments. For RV lateral, in order to utilize its RV context, we constructed the

context parameters as: positions $\{(x_{La} + x_{po})/2, (y_{La} + y_{po})/2\}$; orientation {orthogonal to the line segment connecting RV Lateral and RV insertion posterior, and pointing toward RV insertion anterior side}; and scales $\{s_{La}, s_{La}\}$, where $s_{La} = \sqrt{((x_{La} - x_{po})^2 + (y_{La} - y_{po})^2)} * \gamma$. γ is set to 2.4 in our experiments.

3 Experiments

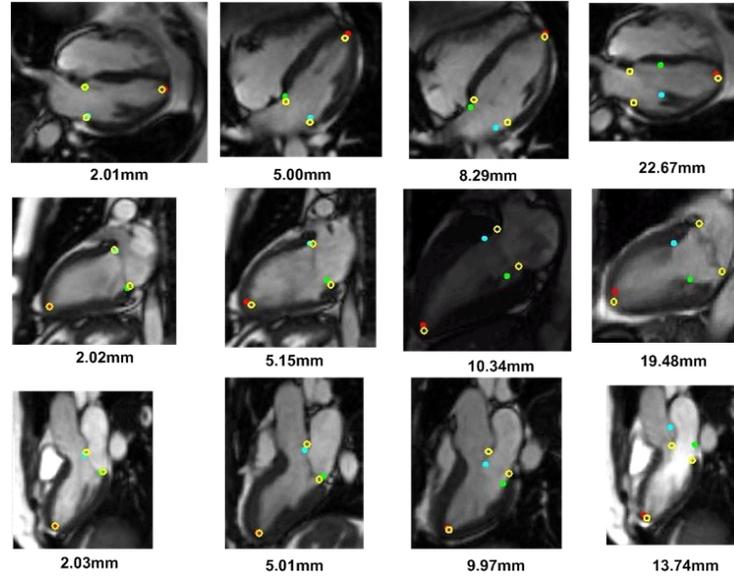
We collected 490 long-axis sequences from 188 patients, whose ages ranged from 11 to 72 years old. In total, 8304 images were used to construct our long-axis image database. Long-axis image can contain different chamber views, namely, 4-chamber, 3-chamber, and 2-chamber views. For each image, the base plane (two annulus anchors) and the apex of the LV were manually annotated by experts and used as ground truth for evaluation. Our short-axis database contains 891 images from 756 sequences of 338 patients, which were provided by two different vendors, each providing 296 and 42 patient studies, respectively. For short-axis images, the ground truth positions of the two RV insertion anchors and the RV lateral anchor were annotated for evaluation purposes.

We applied our context learning algorithm to detect the objects of ‘base plane’, ‘apex’, ‘RV insertion’, and ‘RV lateral’ on respective long-axis and short-axis images for anchoring purposes. The long-axis and short-axis images can be distinguished using orientation information captured during acquisition. With each object detected, corresponding anchors are inferred based on reverse object/context modeling process, i.e., calculating anchor positions from the detected parameterized bounding box. We computed Euclidean distance between the detected anchor position and its corresponding ground truth as the detection error for each anchor. The average distance of all anchors in each image was used as the metric to evaluate the overall system performance.

A 4-fold cross-validation scheme was applied for evaluation. The entire database was randomly partitioned into four quarters. For each fold evaluation, three quarters were combined for training and the remaining one quarter was used as unseen data for testing. This procedure was repeated four times so that each image was used once for testing. Performance is summarized based on all 4 folds and provided in Table 1 and Figs. 4(a) and 4(b). Fig. 3 shows examples of the detection results along with the calculated distance metrics to provide a visual correlation between the distance and quality of detection. In addition to the large parameter search space, cardiac MR images in a large population present a large variation of appearance intensities along with the anatomy shape changes across the heart beat cycle, leading to difficulties for accurate identification. The significant performance difference between RV insertion and RV lateral is due to lack of consistent definition of RV lateral. On the average, it took about 1.5 seconds to detect the base plane and the apex on a 400×400 long-axis image on a duo core 2.8GHz CPU, and 0.5 seconds to detect RV insertion and RV lateral anchors on a 256×256 short-axis image.

Table 1. Average distance of all detected anchors from ground truth positions by a 4-fold cross validation. Distances are in unit of *mm*.

(a) Long-axis				(b) Short-axis			
	Mean	Std	Median		Mean	Std	Median
Overall	4.9	7.0	3.8	Overall	6.7	15.9	4.8
Baseplane	5.1	6.8	3.7	RV insertion	5.9	16.0	3.9
Apex	4.5	6.7	3.3	RV lateral	8.4	16.5	5.9

**Fig. 3.** Examples of anchoring results along with annotated ground truth (yellow). The average detection distance is provided below each image. Our mean and median distances are 5.5mm and 3.8mm, respectively.

In addition, we were able to conduct an independent evaluation on another large expert validated database, which is collected as a mixture of multiple vendors and called ONTARGET database. Patients with vascular disease at high risk of cardiac events were imaged as part of the ONTARGET MRI substudy, described in [14]. Standardized cardiac MRI exams were performed in six countries around the world using Siemens, Philips and GE scanners. Either prospectively or retrospectively gated steady state free precession (SSFP) CMR cines were acquired in six equally spaced short axis (SA) locations from apex to base. Typical imaging parameters were TR/TE/flip/FOV = 30ms/1.6ms/60°/360mm, slice thickness 6mm, image matrix 256×208. There were typically 25 temporal cine frames per slice, depending on the heart rate. All cines were acquired during breath-holding of 815 seconds duration. Most patients had coronary heart disease (87%), myocardial infarction (58%), and/or hypertension (61%). 23% were female and 36% were Asian. The entire database contains 7140 long-axis images

from 87 patient studies. Only base plane anchoring is evaluated. The experimental results are presented in Table 2 and Fig. 4(c). A performance breakdown based on the chamber views is reported in Fig. 4(c). For this evaluation, all original 8304 images were used for training. Then the system was evaluated on the independent 7104 images. Overall, the independent evaluation results are consistent with the internal cross-validation results, both showing promise and robustness of the proposed approach.

Table 2. Average distance of the detected two base plane anchors from ground truth positions from the independent evaluation. Distances are in unit of *mm*.

	Mean	Std	Median
Long-axis	5.2	7.8	3.9

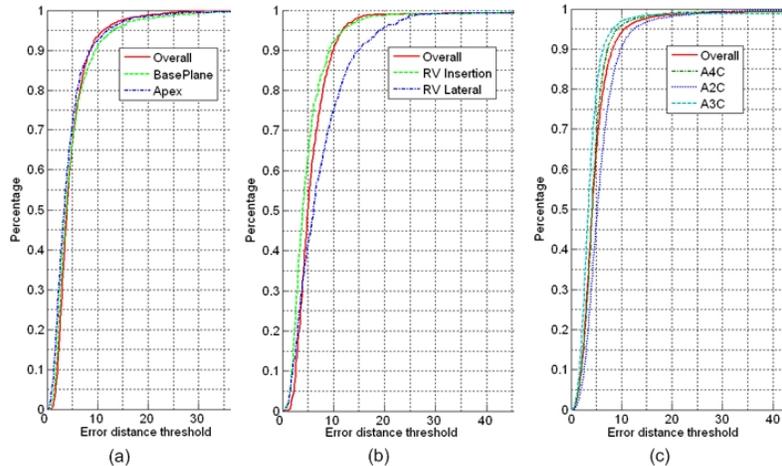


Fig. 4. Percentile evaluation results of 4-fold cross-validation. Each curve shows the percentage of cases whose distance of the automatic detection results from ground truth is less than an error distance threshold. (a) Long-axis anchoring evaluation. (b) Short-axis anchoring evaluation. (c) ONTARGET baseplane evaluation along with performance breakdown based on the chamber views.

4 Conclusions

We have proposed a unified approach to explore contextual information and integrated it with a learning-based object detection framework. We have developed a fully automatic system for cardiac anchoring in both MR long-axis and short-axis images. The principle of the proposed approach is generic. With concrete

design, which is application-specific, the proposed approach is able to be applied to a wide range of applications in addition to acquisition planning and cardiac segmentation.

References

1. Cousty, J., Najman, L., Couprie, M., Clement-Guinaudeau, S., Goissen, T., Garot, J.: Automated, accurate and fast segmentation of 4D cardiac MR images. In: Proc. Functional Imaging and Modeling of the Heart, LNCS 4466. (2007) 474–483
2. Frangi, A., Niessen, W., Viergever, M.: Three-dimensional modeling for functional analysis of cardiac images: A review. *IEEE Trans. on Medical Imaging* **20**(1) (2001) 2–25
3. Finn, J.P., Nael, K., Deshpande, V., Ratib, O., Laub, G.: Cardiac MR imaging: State of the technology. *Radiology* **241**(2) (2006) 338–354
4. Koikkalainen, J., Pollari, M., Lotjonen, J., Kivisto, S., Lauerma, K.: Segmentation of cardiac structures simultaneously from short- and long-axis MR images. In: Proc. MICCAI. (2004) 427–434
5. Jolly, M.P.: Automatic segmentation of the left ventricle in cardiac MR and CT images. *International Journal of Computer Vision* **70**(2) (2006) 151–163
6. Young, A., Cowan, B., Thrupp, S., Hedley, W., Dell’Italia, L.: Left ventricular mass and volume: Fast calculation with guide-point modeling on MR images. *Radiology* **216**(2) (2000) 597–602
7. Weaver, A., Jones, H., Jelaco, G., Cha, J., Hall, J., Walker, K., Blatter, D., Anderson, J.: Magnitude and causes of interobserver discrepancies in CMR volume measurements: critical importance of choice of the basal slice. *Journal of Cardiovascular Magnetic Resonance* **8** (2006) 82–83
8. Cerqueira, M., Weissman, N., Dilsizian, V., Jacobs, A., Kaul, S., Laskey, W., Pennell, D., Rumberger, J., Ryan, T., Verani, M.: Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart. *Circulation* **2002** (2002) 105–539
9. Viola, P., Jones, M.J.: Robust real-time face detection. *International Journal of Computer Vision* **57**(2) (2004) 137–154
10. Georgescu, B., Zhou, X., Comaniciu, D., Gupta, A.: Database-guided segmentation of anatomical structures with complex appearance. In: Proc. IEEE CVPR. (2005)
11. Tu, Z.: Probabilistic boosting-tree: Learning discriminative models for classification, recognition, and clustering. In: Proc. ICCV. (2005) 1589–1596
12. Zheng, Y., Barbu, A., Georgescu, B., Scheuering, M., Comaniciu, D.: Fast automatic heart chamber segmentation from 3D CT data using marginal space learning and steerable features. In: Proc. ICCV. (2007)
13. Lu, X., Georgescu, B., Littmann, A., Mueller, E., Comaniciu, D.: Discriminative joint context for automatic landmark set detection from a single cardiac MR long axis slice. In: Functional Imaging and Modeling of the Heart (FIMH), LNCS 5528, Nice, France (2009) 457–465
14. Cowan, B., Young, A., Anderson, C., Doughty, R., Krittayaphong, R., Lonn, E., Marwick, T., Reid, C., Sanderson, J., Schmieder, R., Teo, K., Wadham, A., Worthley, S., Yu, C., Yusuf, S., Jennings, G.: Left ventricular mass and volume with telmisartan, ramipril, or combination in patients with previous atherosclerotic events or with diabetes mellitus (from the ONgoing Telmisartan Alone and in combination with Ramipril Global Endpoint Trial [ontarget]). *American Journal of Cardiology* (In press, 2010)