

Cross-modality Assessment and Planning for Pulmonary Trunk Treatment using CT and MRI imaging

Dime Vitanovski^{1,2}, Alexey Tsymbal¹, Razvan Ioan Ionasec¹, Bogdan Georgescu¹, Martin Huber¹, Andrew Taylor³, Silvia Schievano³, Shaohua Kevin Zhou¹, Joachim Hornegger², and Dorin Comaniciu¹

¹ Integrated Data Systems, Siemens Corporate Research, Princeton, USA

² Pattern Recognition Lab, Friedrich-Alexander-University Erlangen-Nuremberg, Germany

³ Great Ormond Street Hospital for Children, London, England

Abstract. Congenital heart defect (CHD) is the primary cause of death in newborns, due to typical complex malformation of the cardiac system. The pulmonary valve and trunk are often affected and require complex clinical management and in most of the cases surgical or interventional treatment. While minimal invasive methods are emerging, non-invasive imaging-based assessment tools become crucial components in clinical settings. For advanced evaluation and therapy planning purposes, cardiac computed tomography (CT) and cardiac magnetic resonance (CMR) are important non-invasive investigation technique with complementary properties. Although, characterized by high-temporal resolution, CMR does not cover the full motion of the pulmonary trunk. The sparse CMR data acquired in this context include only one 3D scan of the whole heart in the end-diastolic phase and two 2D planes (long and short axes) over the whole cardiac cycle. In this paper we present a cross-modality framework for the evaluation of the pulmonary trunk, which combines the advantages of both, cardiac CT and CMR. A patient-specific model is estimated from both modalities using hierarchical learning-based techniques. The pulmonary trunk model is exploited within a novel dynamic regression-based reconstruction to infer the incomplete CMR temporal information. Extensive experiments performed on 72 cardiac CT and 74 MR sequences demonstrated the average speed of 110 seconds and accuracy of 1.4mm for the proposed approach. To the best of our knowledge this is the first dynamic model of the pulmonary trunk and right ventricle outflow track estimated from sparse 4D MRI data.

1 Introduction

Congenital heart defect (CHD) is the primary cause of death in newborns characterized by complex malformations of the heart and great vessels. Often, the right side of the heart is affected and especially the pulmonary trunk, as in Tertalogy of Fallot (TOF) and pulmonary artesia or stenosis. The clinical management of such conditions is confronted with complex treatment decisions, which include pulmonary valve procedures in the majority of the cases.

Percutaneous interventions for pulmonary valve replacement are emerging as feasible treatment alternatives to classic cardiac surgery with important benefits: less invasive, reduced risks associated with cardiopulmonary bypass, bleeding, infections and reduced expenses for postoperative intensive care [1]. Nevertheless, comprehensive investigation, based on non-invasive imaging techniques, is mandatory for clinical decision making and treatment success.

For therapy planning purposes, the pulmonary trunk is increasingly imaged using either cardiac computer tomography (CT) or cardiac magnetic resonance (CMR) [2]. While CT has a high spatial resolution, fast acquisition times without anesthesia, it has the disadvantages of poor temporal resolution and ionizing radiation. Contrary, MRI has high temporal resolution without X-ray radiation, but long acquisition times and usually does not cover the full 4D information. The regular protocol, so called sparse 4D CMR, involves an end-diastolic (ED) 3D heart image and two orthogonal cine projections 2D+t, short axis (SA) and long axis (LA). LA passes through the main pulmonary artery and the descending aorta, while SA is aligned with pulmonary valve, perpendicular to the LA (see Fig. 1). Ideally, clinicians would be provided with accurate morphological and functional quantification of the pulmonary trunk, independent of the employed imaging technique.

In this paper we present a cross-modality framework for the evaluation of the pulmonary trunk, which combines the advantages of both, cardiac CT and CMR, non-invasive imaging techniques. A physiological model, which captures complex morphological, dynamical and pathologic variations of the pulmonary trunk is presented in Sec. 2. In Sec. 4, the patient-specific model parameters are estimated from both modalities within hierarchical learning-based framework, which involves three-stages: landmark detection, center line detection and dynamics estimation. A novel dynamic regression-based reconstruction is proposed to infer the incomplete temporal information characteristic to the sparse MRI protocols.

Extensive experiments are performed on 72 cardiac 4D CT (720 volumes) and 74 sparse MRI (74 3D ED volumes associated with 4736 2D slides over the cardiac cycle) data, from which 10 patients underwent both imaging interventions, CT and sparse MRI. Mean reconstruction error of 1.44 mm within 110 seconds demonstrates the strength of our proposed regression based reconstruction method.

2 Physiological Pulmonary Trunk Modelling

In this section we introduce our physiological model of the RVOT and pulmonary trunk, which represents both morphological and dynamical variations. Similar as in [3], the anatomical complexity is reduced by employing a coarse to fine parameterization which includes: anatomical landmarks, pulmonary artery center line and full surface model of the pulmonary trunk. As illustrated in Fig. 1(a), the considered anatomical landmarks include, Trigone (L_t), RVOT (L_{rvot}) and Main-Bifucation (L_{mb}), each represented in the Euclidean 3D space, $L_x \in R^3$. The centre line CL passes through the pulmonary artery center and is parameterized by 12 points, $CL = CL_0 \dots CL_{11}$. The surface model S is represented by

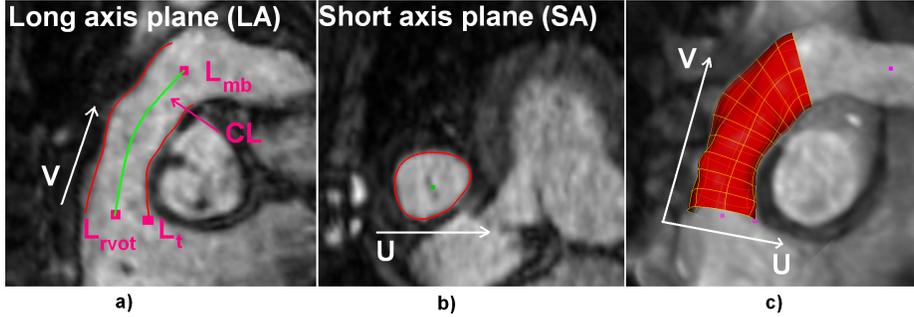


Fig. 1. 3D MRI scan of the whole heart in the ED phase (a). 2D long axis (LA) plane (b) and short axis (SA) plane (c) of the pulmonary artery over the cardiac cycle

a structured grid, spanned along two anatomical directions, u - *circumferential* and v - *longitudinal*, using 50×40 vertices (see Fig. 1(c)). Point correspondence in time and across patients is enforced by intrinsic re-sampling of S , using a set of anatomical-driven cutting-planes, described by center line points CL_x and corresponding tangential directions. Given the different modalities supported and characteristic imaging protocols, we differentiate among two dynamic extensions of the proposed physiological model. The definition of a full 4D model, which can be directly estimated from 4D cardiac CT data, is rather straightforward and realized by concatenating a time variable t :

$$Model_{full4D} = \{L_t, L_{rvot}, L_{mb}, CL, S\}_t \quad (1)$$

However, given the sparse 4D acquisition, common to MRI exams, the extension to a temporal model includes two additional representations: LA and SA . LA - describes the contour of S intersected with the plane with the origin in L_{rvot} and the normal obtained from the cross-product between the $L_{trigone}$ and L_{rvot} , and the center line tangent at CL_0 - these represents a specific 2D+time long axis acquisition (Fig. 1(a)). SA - describes the contour of S intersected with the plane center in CL_{middle} and the corresponding tangent as normal (Fig. 1(b)). Hence, the sparse dynamic model is parameterized as follows:

$$Model_{sparse4D} = \{L_t, L_{rvot}, L_{mb}, CL, S\}_{ED} + \{LA, SA\}_t \quad (2)$$

3 Dynamic Regression Based 4D Model Reconstruction

As describe above, the dynamic information in case of MRI 4D is incomplete but rather available only in two orthogonal projections, as opposed to 4D cardiac CT, which provides full dynamics over the cardiac cycle. In this section we propose a novel dynamic regression method to learn the pulmonary trunk spatial-temporal variation from a heterogonous training population and apply it to predict a full dynamic model from partial 4D usually provide by MRI.

In *regression* a solution to the following optimization problem is normally sought [4]:

$$\hat{\mathcal{R}}(\mathbf{x}) = \underset{\mathcal{R} \in \mathfrak{S}}{\operatorname{argmin}} \sum_{n=1}^N L(y(\mathbf{x}_n), \mathcal{R}(\mathbf{x}_n)) / N \quad (3)$$

where \mathfrak{S} is the set of possible regression functions, $L(\circ, \circ)$ is a loss function that penalizes the deviation of the regressor output $\mathcal{R}(\mathbf{x}_n)$ from the true output, and N is the number of available training examples. In our case the reconstruction task is defined as a regression problem between the full dynamic model of the pulmonary trunk extracted from 4D CT data and the sparse one extracted from the sparse MRI data:

$$Y(\operatorname{Model}_{full4D}) = \hat{\mathcal{R}}(X(\operatorname{Model}_{sparse4D})) + \epsilon \quad (4)$$

In our regression problem we focus on shape information and completely neglect volume data. As descriptors both for the input $X(\operatorname{Model}_{sparse4D})$ and output elements $Y(\operatorname{Model}_{full4D})$ of the models we choose *coordinates* of mesh vertices normalized with the generalized procrustes analysis. This representation has been already used before, with the purpose of model classification into diseased and healthy, and has a uniform representation of the input and the output data. The training set T used to generate the regression model includes feature vectors T_i as follows:

$$T_i = \langle (S_i^{ED}, LA_i^t, SA_i^t)_{MRI}, (S_i^t)_{CT} \rangle, \quad (5)$$

where t is the time step within the cardiac cycle, S_i^{ED} is a set of 3D coordinates representing each point of the end-diastolic model (2000 3D points), LA_i^t and SA_i^t are point sets (80 and 50 3D points respectively) representing the model curves extracted from the MRI's long axis stack and short axis stack respectively, for the current time step t , and $(S_i^t)_{CT}$ are the corresponding point coordinates for the point set to be reconstructed (238 3D points). Due to the dense representation of our model (2000 3D points) we reconstruct only the most significant 238 3D points from the associated CT model. The rest of the points are interpolated and projected onto the PCA shape space from which the complete final model is then obtained.

The formulated regression problem is solved by learning the regression function \mathcal{R} with two different methods: *boosting-based additive regression* [5] and *random forest* [6]. Two main reasons motivate our choice. First, these techniques were shown to be robust to high-dimensional data with many irrelevant, redundant and noisy features, without the need for additional data pre-processing and feature selection. This was shown both for classification [7],[8] and regression [4] tasks. Second, both boosting-based and random forest-based models are relatively fast to train and to evaluate comparing for example with Support Vector Regression. In the spirit of [7],[4], we use *simple 1D linear regression* as the base learner for boosting-based regression. At each boosting iteration, a feature which results in the smallest squared loss with linear regression is added to the pool of already selected features. Each weak learner is thus a simple linear regressor of the form:

$$y = \beta_0 x + \beta_1 \quad (6)$$

where x is the selected scalar input coordinate and y is a scalar output coordinate. Using more sophisticated weak learners such as CART decision trees and multiple linear regression with greedy forward feature inclusion, has proven to always result in a worse or no better performance while the resulting model gets significantly more complicated. Using simple 1D binary decision stumps as in [4] has also proven to lead to suboptimal accuracy; the reason for this is perhaps the nature of the data, as it is rather impossible to generate as many candidate decision stumps with the coordinate - based features as it is possible with the Haar-like features. For each boosting-based model, we generate 200 weak learners. The accuracy plateaus with this number of component models, and the further accuracy increase is always insignificant with this data.

For random forests, we always generate 25 component trees. The accuracy usually remains same or even decreases with the addition of more trees to the model. The minimum leaf size is set to 1; the trees are thus generated to the full with no pruning. The number of features considered at each node is set to the value recommended by Breiman [6], which is one third of the total number of features for regression. Using other parameter settings was shown to lead to worse or no better accuracy in our preliminary experiments.

In boosting-based regression the output function is assumed to take a linear form as follows [4]:

$$\hat{\mathcal{R}}(\mathbf{x}) = \sum_{t=1}^T \alpha_t h_t(x) \in H \quad (7)$$

where $h_t(x)$ is a base (weak) learner and T is the number of boosting iterations. Having a *linear* base learner (simple linear regression), a *linear* final solution is thus also found. In contrast to this, random forests seek for a non-linear function approximation, recursively splitting the feature space in the nodes of component decision trees.

In contrast to [4], we use *naive decoupling* of the regression problem into a number of single output problems. While multi-output regression solutions do exist both for boosting [4], for our task multi-output optimization was not shown to lead to error decrease and time savings were rather insignificant.

4 Estimating Patient-Specific Model Parameters

The patient-specific model parameters described in Section 2 are estimated from cardiac acquisition using a learning-based algorithm. Detectors are learned separately for both modalities, CT and MRI, and applied to estimate model parameters in a hierarchical three-stage approach: Anatomical Landmarks Estimation, Center Line Estimation and Full Surface Model Estimation.

Anatomical Landmarks Estimation By defining the localization as a classification problem, the anatomical landmarks, L_t, L_{rvot}, L_{mb} , are estimated within the Marginal Space Learning (MSL) framework [9]. Separate detectors $D_t^L, D_{rvot}^L, D_{mb}^L$, are learned using the Probabilistic Boosting Tree (PBT) [8] in combination with Haar-like feature from a training dataset annotated by experts.

$$p(L_x|x, y, z) = D_x^L(x, y, z), (x, y, z) \in \sigma^x \quad (8)$$

the trained detectors D_x^L models the target posteriori distribution $p(L_x|x, y, z)$ for a specific search space σ^x given by the training set. MSL is applied to exhaustively search the parameter space using the learned detectors and obtained the location of the anatomical landmarks. Note that in case of 4D cardiac CT anatomical landmark are detected in each volume to obtain the dynamic parameters $\{L_t, L_{rvot}, L_{mb}\}_t$, while in sparse cardiac MR only a static detection in the end-diastolic volume is performed $\{L_t, L_{rvot}, L_{mb}\}_{ED}$,

Center Line Estimation CL passes through the centre of the pulmonary trunk and is initialized by the previously detected landmarks L_t and L_{rvot} . A robust detector D^{CL} is learned using the same MSL framework to detect circular structures, parameterized by center line points CL_x , corresponding tangent and fixed radius $r = 20mm$ obtained from the average value in the training set. An incremental approach is used to search circles on a series of successively updating planes. Please note, as for the Anatomical Landmarks Estimation, a temporal center-line model CL_t is detected in CT and a static CL in MR.

Full Surface Model Estimation The full model of the pulmonary trunk S is initialized in the end-diastolic frame using the estimated landmarks and centerlines, using a piecewise affine transformation along the center line [3]. Robust boundary detectors D^s , trained using the PBT and steerable feature [9] are applied to locally refine the surface by moving it along normal directions towards the position with highest boundary probability. To obtain spatially smooth delineation, the final results is obtained by projecting S to a previously learned shape space model.

In case of CT, the above describe algorithm is applied in each time step to obtain the full temporal model $\{L_t, L_{rvot}, L_{mb}, CL, S\}_t$. In case of MR, the estimated surface in the end-diastolic frame S_{ED} is used to initialize the contours LA and SA . These are refined using a trained D^c contour detector as described above. A full dynamic 4D model is then estimated by using a learned regression model (see Eq. 4) to predict the missing temporal information.

5 Results

5.1 Results on Patient-Specific Model Parameters Estimation

The proposed framework for detecting a personalized pulmonary trunk model in 4D CT and sparse MRI data was evaluated on 50 4D CT(500 volumes) and 74 sparse MRI (74 ED Volumes associated with 4736 LA/SA planes) studies from patients with different CHD. Each volume in the data set is associated with annotation, manually generated by experts, which is considered as ground truth. Three-fold cross validation was used to divide the data set into training and test data.

Table 1 summarizes the detection performance on both modalities (CT and sparse MRI), from the test data. Point-to-mesh measurement error was used to evaluate the detection accuracy between the ground-truth and detected model for both modalities. Average speed of 10sec per frame was achieved for both modalities on a standard 2.0GHz Dual Core PC.

Table 1. Detection accuracy

CT/MRI	Mean Error(mm)	Median(mm)	Std.Dev(mm)
Landmarks	3.5/4.3	5.1/6.4	2.7/3.0
Center Line	3.0/3.3	2.3/2.3	1.7/2.0
Full Surface	1.6/1.9	1.2/1.3	0.2/0.2

5.2 Intra-modality comparison between CT and MRI

The inter-modality consistency of the model was demonstrated on a subset of 10 patients which underwent both imaging investigations, 4D CT and sparse MRI (see Fig. 2). Ground-truth and detected pulmonary trunk models from both modalities were compared using the abstract point-to-mesh measurement and clinical relevant diameter measurements: RVOT, hinges and commissures. Results are summarized in Table 2. A strong inter-modality correlation, $r = 0.992$, $p < 0.0001$ and confidence of 98%, was obtained for CT and CMR based on the pulmonary trunk model.

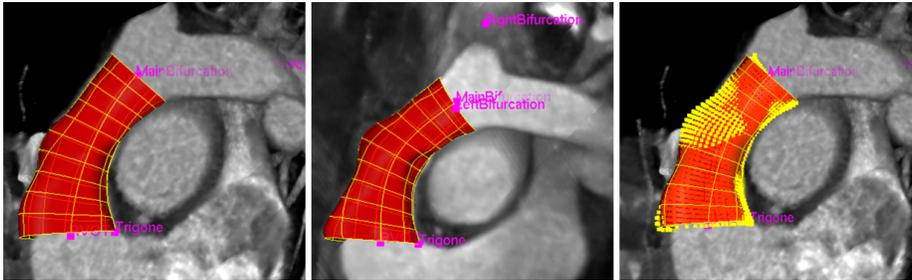


Fig. 2. Pulmonary trunk model in CT (**left**) and MRI (**middle**) data for the ED phase. Inter-modality consistency by projecting the MRI model(yellow points) into the CT data (**right**).

Table 2. Model based intra-modality comparison between CT and MRI

(mm)	Ground truth	Estimation
RVOT	0.7 ± 0.5	3.8 ± 1.5
Hinges	1.2 ± 1.4	2.6 ± 4.7
Commissures	1.5 ± 1.2	3.2 ± 1.7
Point-to-mesh	1.4 ± 0.1	2.5 ± 0.7

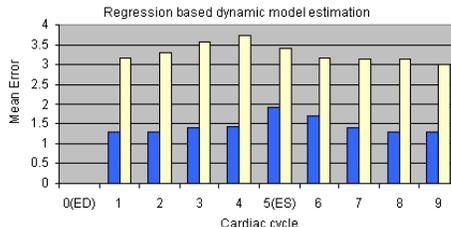
5.3 Results on Regression Based Dynamic Model Reconstruction

As described in Section 1 the sparse MRI protocol is able to capture the full anatomy of the pulmonary trunk only in the ED phase (3D volume) of the heart and parts of the pulmonary trunk in 2D planes (LA and SA) over the cardiac cycle. However, a full 4D model of the pulmonary trunk can be still computed from the available sparse data by learning the full motion from 4D CT data. For this purpose we learned a regression model as presented in Section 3 on a

training data set of 72 4D CT (720 Volumes) studies. Two different machine-learning techniques (boosting and random forest) are used to train the regression model and to evaluate the reconstruction error. Table 3 presents values obtained by applying the regression method on sparse CMR images and evaluate it on full 4D CT, for a set of 10 patients, which underwent both imaging modalities. Figure 3 illustrates the reconstruction error distributed over the cardiac cycle.

Table 3. Reconstruction error for Random Forest and Boosting

	Boosting	Random Forest
Mean Err.	1.44(mm)	3.2(mm)
Std. Dev	0.21(mm)	0.23(mm)
Speed	3.07 (ms)	6.21 ms



6 Conclusion

In this paper we propose a cross-modality detection framework for estimating a dynamic personalized model of the pulmonary trunk from the available data, 4D CT and sparse MRI. A novel regression based reconstruction method is presented and used to infer the incomplete temporal information characteristic to the sparse MRI protocols. The estimated model from both modalities can be utilized to extract morphological and functional information of the pulmonary trunk and dynamics over the cardiac cycle. Extensive experiments performed on a large heterogeneous data set demonstrated a precision of 1.44mm data at a speed of 11 seconds per volume. The proposed method has the potential to significantly advance the pulmonary trunk treatment.

References

1. Bonhoeffer, P.e.a.: Percutaneous insertion of the pulmonary valve. *Journal of the American College of Cardiology* **39**(10) (May 2002) 1664–1669
2. Taylor, A.: Cardiac imaging: Mr or ct? which to use when. *Pediatr Radiology* **38** (2008)
3. Vitanovski, D., Ionasec, R., Georgescu, B., Huber, M., Taylor, A., Hornegger, J., Comaniciu, D.: Personalized pulmonary trunk modeling for intervention planning and valve assessment estimated from CT data. In: *MICCAI*. (2009) 17–25
4. Zhou, S.K., Georgescu, B., Zhou, X.S., Comaniciu, D.: Image based regression using boosting method. In: *ICCV*. (2005)
5. Friedman, J.H.: Greedy function approximation: A gradient boosting machine. *Annals of Statistics* **29** (2000) 1189–1232
6. Breiman, L.: Random forests. In: *Machine Learning*. (2001) 5–32
7. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: *CVPR*. (2001)

8. Tu, Z.: Probabilistic boosting-tree: Learning discriminative methods for classification, recognition, and clustering. In: ICCV. (2005)
9. Zheng, Y., Barbu, A., et al.: Fast automatic heart chamber segmentation from 3d ct data using marginal space learning and steerable features. ICCV (2007)