Learning discriminative distance functions for valve retrieval and improved decision support in valvular heart disease

Ingmar Voigt\textsuperscript{a,c}, Dine Vitanovski\textsuperscript{a}, Razvan Ioan Ionasec\textsuperscript{b,d}, Alexey Tsymbal\textsuperscript{a}, Bogdan Georgescu\textsuperscript{b}, S. Kevin Zhou\textsuperscript{b}, Martin Huber\textsuperscript{a}, Nassir Navab\textsuperscript{d}, Joachim Hornegger\textsuperscript{c} and Dorin Comaniciu\textsuperscript{b}

\textsuperscript{a} Siemens Corporate Technology, SE SCR 2, Erlangen, Germany
\textsuperscript{b} Integrated Data Systems, Siemens Corporate Research, Princeton, USA
\textsuperscript{c} Chair of Pattern Recognition, University of Erlangen-Nuremberg, Germany
\textsuperscript{d} Computer Aided Medical Procedures, Technical University Munich, Germany

\textbf{ABSTRACT}

Disorders of the heart valves constitute a considerable health problem and often require surgical intervention. Recently various approaches were published seeking to overcome the shortcomings of current clinical practice, that still relies on manually performed measurements for performance assessment. Clinical decisions are still based on generic information from clinical guidelines and publications and personal experience of clinicians. We present a framework for retrieval and decision support using learning based discriminative distance functions and visualization of patient similarity with relative neighborhood graphs based on shape and derived features. We considered two learning based techniques, namely learning from equivalence constraints and the intrinsic Random Forest distance. The generic approach enables for learning arbitrary user-defined concepts of similarity depending on the application. This is demonstrated with the proposed applications, including automated diagnosis and interventional suitability classification, where classification rates of up to 88.9\% and 85.9\% could be observed on a set of valve models from 288 and 102 patients respectively.

\textbf{Keywords:} physiological valve modeling, machine learning, discriminative distance function, case retrieval, decision support

\section{1. DESCRIPTION OF PURPOSE}

Valvular heart disease (VHD) is an important subgroup of the cardiovascular diseases, which is known to be the No. 1 lethal disease worldwide.\textsuperscript{1} Among the latter, treatment of VHD is most expensive and has the highest in-hospital death rate, due to elaborate, time consuming and potentially inaccurate diagnostic procedures and complex interventions into one of the most essential physiological systems of the human body. Recent advances in scanner technology enable for 4D imaging with CT and Ultrasound. These modalities are well suited to the non-invasive capture of valve morphology and dynamics. Due to the lack of efficient and convenient tools however anatomical performance assessment still relies on manual measurements in 2D image planes, derived from the acquisitions, which is error prone and time consuming. Diagnosis, treatment decisions, interventional planning and follow up evaluation however rely on such performance assessment and may lead to suboptimal treatment results, follow up interventions and increased costs.

Case-based reasoning (CBR) systems are an important subgroup of clinical decision support systems and provide a way to retrieve similar cases and related clinical data such as diagnoses, treatment decisions, follow up results and long term outcomes. Important factors for a well usable CBR however are: defining a meaningful distance function, which needs to tackle the complexity of clinical data, and a transparent way to visualize the underlying inter-patient similarity, which is the central concept of any clinical CBR.

In this paper we present a generic method on how to automatically derive high-level information from geometric valve models and derived features using learning based discriminative distance functions. We use relative neighborhood graphs\textsuperscript{2,3} to visualize the relationships among the cases in an intuitive way. Our method is generic.
Figure 1. The proposed models of the aortic-mitral complex (a) and pulmonary trunk (b) in from the same CT data set and all valves in visualized together (c), showing their spatial relation.

and flexible and enables for arbitrarily defining and learning similarity among shapes based on any meaningful concept. It is applied to physiological models of heart valves, which cover their complete morphology and are modeled from 4D cardiac CT, 4D transesophageal Ultrasound and 3D isotropic MRI respectively. Finally we present a set of clinically relevant applications for the management of valvular heart disease including decision support and retrieval, which may have significant impact on health care and generates fully reproducible results.

The remaining paper is structured as follows: Section 2.1 briefly describes the models and how they are derived from the acquisitions, while Section 2.2 gives an overview on learning based distance functions. Section 2.3 describes the technique of neighborhood graphs for patient similarity visualization. The clinical applications are then summarized in section 2.4; the results are then reviewed and the paper concludes with Sections 3 and 4.

2. METHODS

2.1 Heart Valve Modeling

The morphology and dynamics of the aortic-mitral complex and pulmonary trunk is described by the 4D physiologically compliant models estimated from CT and Ultrasound, as proposed by Ionasec et al.\(^5\) and Vitanovski et al.\(^6\) The central anatomical structures are the aortic root and leaflets, the mitral leaflets and the pulmonary trunk, represented by Non-uniform rational B-Splines (NURBS), and a set of clinically relevant anatomical landmarks, which constrain them topologically and geometrically to capture the large morphological and pathological variation.

The model estimation works in an automatic and hierarchical fashion by applying robust discriminative learning and incremental searching techniques at each step,\(^5,6\) allowing for manual correction if desired by the user. These steps include estimation of the piecewise affine rigid transformation, landmark detection, followed by centerline detection for the pulmonary trunk, and finally boundary fitting. For each stage \(i\) a discriminative classifier \(H_i\) was trained using the Probabilistic Boosting Tree\(^7\) in combination with Haar-like and steerable features.\(^8\) \(H_i\) is then used to incrementally scan the search space of acquisition \(I\) to determine the parameters \(\hat{\theta}_i\) with maximum probability:

\[
\hat{\theta}_i = \arg\max_{\theta_i \in \Theta_i} p(\theta_i|I) = H_i(\theta_i|I)
\] (1)

The results of each stage are propagated to subsequent stages in order to initialize and constrain the search spaces of the successive estimation problems, by piecewise rigid and non-rigid transformations respectively. To
tackle the problem of temporally varying appearances and image characteristics we exploit the temporal coherence by using incrementally learning and estimating landmark trajectories in the Fourier space and ISOMAP motion models to predict the shape deformation over time. Fig. 1 depicts models of the aortic-mitral complex and the pulmonary trunk.

Clinically established measurements, characterizing the valve’s morphology and the severity of the various diseases, are derived from the models in a straightforward manner by fast and simple but efficient geometrical computations such as diameters, annular and valvular areas, inter-landmark distances and angles based on clinical observations and research. Fig. 2 shows our valve models together with a selection of measurements.

In our experiments on distance learning and retrieval, the shape $S$ of the valve models is represented by a discrete set of $N$ 3D points $P_i$ sampled from the NURBS surfaces $S_i$:

$$S = \{ P_i = (P_i^x, P_i^y, P_i^z) \mid P_i = S_i(u, v); u, v \in [0; 1]\} \quad (2)$$

In order to achieve an accurate representation, 951 discrete points are sampled from the aortic valve surfaces altogether, while the pulmonary trunk requires representation by 2000 points.

### 2.2 Learning discriminative distance functions

Generally we address two tasks: retrieval of similar cases using a learned distance function, which measures the similarity of two particular shapes, and a binary classification problem, based on geometric valve models and derived features as described in the previous section. The second task can actually be formulated using the first.\textsuperscript{2,4} So instead of learning a classifier directly, we choose learning distance functions instead and divide the learning process into two sequential steps, i.e. distance learning followed by classification or clustering, where each step requires search in a less complex functional space than in the immediate learning.\textsuperscript{4}

The labels for classification $y \in \{-1, +1\}$ are chosen depending on the application. Each case is represented with a parameter vector $C$ containing the $N$ 3D points $P_i$ of the respective model and a set of $M$ additional features $F_i$ derived from the model, depending on the application:

$$\hat{y} = \underset{y \in \{-1, +1\}}{\text{argmax}} \ (p(y|C)) \quad C = ((P_1^x, P_1^y, P_1^z), \cdots,(P_N^x, P_N^y, P_N^z), F_1, \cdots, F_M) \quad (3)$$

Learning a distance function helps to combine the power of strong learners with the transparency of nearest neighbor classification.\textsuperscript{9} Moreover, learning a proper distance function was shown to be especially helpful for
high dimensional data with many correlated, weakly relevant and irrelevant features, where most traditional
 techniques would fail. Also, it is easy to show that choosing an optimal distance function makes classifier
 learning redundant.

2.2.1 Equivalence constraints

Learning from weak representations such as equivalence constraints usually provides high flexibility for learning
 arbitrary functions. Equivalence constraints are represented using triplets \((C^1, C^2, y)\), where \(C^1\) and \(C^2\) are
 coordinates in the original space (see equation 3) and \(y \in \{+1, -1\}\) is a label indicating whether the two points
 are similar (from the same class or cluster) or dissimilar.\(^{10}\)

Learning from these triplets is often called learning in the product space (i.e. with pairs of points as input).\(^{9–11}\) Another common alternative is to learn in the difference space, the space of vector differences.\(^{12}\) While both
 representations demonstrate promising empirical results in different contexts, there is no understanding which
 representation is better and when.

Commonly a binary distance function is the output from learning from equivalence constraints, predicting
 only whether the two objects are similar or dissimilar. This limitation can however be overcome by using the
 signed margin of margin-based classifiers such as SVM and boosting as the required distance function.

Clearly, the most popular learning algorithm in the area of distance function learning from weak representa-
tions is boosting.\(^{13}\) We also used Random Forests, which demonstrated promising results on public benchmark
datasets and are normally faster than boosting. Both product and difference spaces were implemented in our
 framework. The signed margin of models constructed using AdaBoost and Random Forests is used as the required
 distance function for our experiments with equivalence constraints.

2.2.2 The intrinsic Random Forest distance function

For a Random Forest (RF) learned for a certain classification problem, the proportion of the trees where two
 instances appear together in the same leaves can be used as a measure of similarity between them.\(^{4,14}\) For a given
 forest \(f\) the similarity between two instances \(C^1\) and \(C^2\) is calculated as follows. The instances are propagated
down all \(K\) trees within \(f\) and their terminal positions \(z\) in each of the trees \((z_1 = (z_{11}, \ldots , z_{1K})\) for \(C^1\), similarly
 \(z_2\) for \(C^2\)) are recorded. The similarity between the two instances then equals to \((I \text{ is the indicator function})\):

\[
S(C^1, C^2) = \frac{1}{K} \sum_{i=1}^{K} I(z_{1i} = z_{2i})
\]  

Compared to learning from equivalence constraints the intrinsic RF distance is rather uncommon in the area
 of learning based distance function and is never considered as a possible alternative. In general, we believe,
 that the circle of applications both for distance learning from equivalence constraints (which is currently applied
 nearly solely to imaging problems) and for the intrinsic RF distance is still, undeservedly, too narrow and may
 and should be expanded.

2.3 Neighborhood graphs for inter patient similarity visualization

Neighborhood graphs provide an intuitive way of patient similarity visualization with an entity-relationship
 representation.\(^2\) In a relative neighborhood graph (RNG), two vertices corresponding to two instances \(C^1\) and
 \(C^2\) in a data set are connected with an edge, if there is no other instance \(C^3\) which is closer to both \(C^1\) and \(C^2\)
 with respect to a certain distance function \(d:\)

\[
\forall C^3 : d(C^1, C^2) \leq \max\{d(C^1, C^3), d(C^2, C^3)\}
\]  

We prefer the RNG for visualization, as an important advantage is, that they are always planar or close to
 planar and connected, while nodes have a reasonable small degree in contrast to related concepts like directed
 nearest neighbor or distance threshold graphs. In comparison to the well known heatmaps, they are easier to
 read with the more intuitive node-link representation, they allow visualizing additional features or even image
 thumbnails at nodes, and they have a flexible layout allowing to naturally visualize clusters, enlarge nodes, and
 filter out a set of nodes and edges.\(^2\)
2.4 Application to heart valves

The basic workflow within our framework is visualized in Fig. 3: A 4D acquisition (Fig. 3(a)) is loaded into the modeling component, which then generates the personalized 4D models (Fig. 3(b)). These are passed on to the retrieval component, which finds similar valves and classifies them depending on the application (PPVI suitability in this case).

A neighborhood graph (Fig. 3(c)) is generated and presented to the user, who can inspect the similar valves and patient records to then find a diagnostic or treatment decision. Such a graph is an expressive way of presenting relevant information. Image thumbnails allow for easily identifying the different cases and their shape characteristics. From the graph, one may easily comprehend patient distribution according to the studied similarity context and see patient groupings, identify outliers, easy to classify cases and the borderline cases classification for which is likely to be uncertain.

In our framework shape similarity can be defined arbitrarily and depending on the application. In the following different applications in field of decision support for VHD are described, which provide example on how shape similarity could be defined and exploited.

2.4.1 Model-driven diagnosis

One application we proposed is the classification of given shapes into diseased and healthy cases. There are two main types of heart valve malfunction: stenosis and regurgitation. Valvular stenosis is defined as an obstruction of the blood flow caused by narrowing, stiffening, thickening, fusion or blockage of one or more valves. Valvular regurgitation is a condition in which blood leaks back in the wrong direction because the affected valve is not closing properly. Fig 4 depicts a selection of pathologic cases in CT and Ultrasound. Both disorders can greatly interfere with the heart’s function to pump an adequate amount of blood, causing serious health problems.

The classification is performed by using a clustered RNG, generated with the underlying learning based distance measure for the input meshes. The resulting graph displays two regions indicating healthy and diseased cases and the mutual similarity among the cases.

Clinically established geometrical measurements usually determine the severity of disease and required interventions and can be easily derived from the models to further enrich the feature pool and strengthen the classification.
2.4.2 Interventional planning in percutaneous pulmonary valve implantation

Pulmonary valve replacement traditionally is done by open heart surgery\textsuperscript{16}, implying risks as stroke and repeated valve replacement\textsuperscript{17}, in contrast to percutaneous pulmonary valve implantation (PPVI)\textsuperscript{18}. The morphology of the pulmonary trunk is a major determinant of suitability for PPVI\textsuperscript{19} as a valve stent placed in patients with pyramidally shaped trunks has a high probability of proximal device dislodgement.

Schievano et al.\textsuperscript{18} proposed classification based on geometric measures and appearance, where the variation of the right-ventricular outflow tract and the pulmonary trunk is classified into five morphological groups (Fig. 5). Patients with a pulmonary trunk morphology of type I as depicted in 5(a) are considered to be unsuitable for PPVI due to the narrow artery and high probability of device migration. Therefore the main challenge lies in discriminating anatomies of type I from other classes. We propose a robust generic shape-based patient selection for PPVI based on learned discriminative distance functions to classify into PPVI suitable and PPVI unsuitable (Fig. 3).

2.4.3 Retrieval for definition of personalized treatment targets

The coupling of the aortic and mitral valvular annuli through fibrous tissue is evident and leads to strong functional and morphological interdependency. Recent studies emphasize their correlation given by the anatomy\textsuperscript{20–22}. We propose to exploit this mutual morphological correlation among the different valves, for retrieval of similar healthy valves for a diseased case, in order to define precise and personalized targets of an intervention.
The following example illustrates the proposed use case: in case of a patient with a diseased mitral valve (MV), affected by annular dilation, and a healthy aortic valve (AV), we retrieve a case with a similar healthy AV and a healthy MV. The latter can then be used to determine the annular dimensions of the formerly healthy MV for mitral annuloplasty, i.e. the system provides a suggestion to the surgeon, which dimension of the prosthesis would actually be the best fit, such that the surgeon has an additional statistically based decision support, rather than having to rely on generic indication from clinical guidelines, publications and personal experience only. Moreover another interesting use case would be the design of personalized prostheses instead of selecting an optimal generic one, which would fit and reflect the formerly healthy valve’s morphology and performance even better.

3. RESULTS

The performance of the proposed method was evaluated on 4D cardiac CT and transesophageal echocardiographic data from 288 patients for the aortic valve diagnosis application (section 2.4.1) and 4D cardiac CT and 3D isotropic MRI data from 102 patients for the pulmonary trunk suitability selection for PPVI (section 2.4.2), with significant variation in image resolution and capture ranges.

For the aortic valve diagnosis experiment the evaluation set includes 234 healthy and 54 diseased valves, where one or multiple attributes of the following could be observed: regurgitation, stenosis, bicuspid malformation, dilation. In total 288 volumes associated with manual annotations where processed. For the PPVI suitability selection experiment the evaluation set includes 50 patients with pulmonary trunk geometry of type I (i.e. unsuitable for PPVI) and 52 patients with suitable geometries, totaling in 102 volumes and associated manual annotations. The population in both datasets covers a great range of ages, as many different disease and appearance patterns are included, such as calcified aortic valves (typically a condition in elderly patients), but also bicuspid aortic valves, which is a congenital malformation (Fig. 4). Each type of pulmonary trunk geometry (Fig. 5) is represented, where many diseased cases are children, while a large portion of healthy valves was modeled from mature patients. Therefore besides pathology and individual variation yet another important factor, which increases the data heterogeneity is patient’s age, influencing the valve morphology as well and complicating the task of learning from such data.

The accuracy of the model based classification is validated by 10-fold cross-validation (Fig.6). The figure demonstrates the classification accuracy of k-Nearest Neighbors (kNN), AdaBoost (AB) and Random Forests (RF) in the canonical space, as well as AdaBoost and Random Forests in the product and difference spaces (AB-pr, RF-pr, AB-di and RF-di) and intrinsic RF distance (RF-dist). Prior to learning the valve shapes were normalized by their individual size and aligned with procrustes analysis, to cancel out translation, orientation
and scale and minimize the effects of age. Each cross-validation run was performed on 200 (aortic) and 600 (PPVI) pre-selected features respectively, which were determined using the Gain-Ratio feature filter.

Learning from equivalence constraints improves the accuracy in comparison with learning in the canonical space. For this task the product space leads to the best accuracy in both experiments. While for the aortic valve experiment RF showed best performance in the product space with a classification rate of 88.9%, AdaBoost in the product space showed higher performance for the PPVI suitability selection with 85.9%. Learning distance functions from equivalence constraints helps to combine the power of boosting and random forests with the transparency of case retrieval. While experimental results in earlier publications\textsuperscript{2,4,6,15} showed better performance, their data sets are incomparably smaller in size and degree of heterogeneity.

For our experiments in the setting of the aortic valve disease classification we investigated the computed Gain-Ratio values for each coordinate of the geometrical model, in order to better understand the correlation between morphology and pathology. The most relevant regions for diseased valves are the free parts of the leaflets (Fig. 7(a)), around the tips, as well as the right coronary ostium region and the upper part of the aortic valve towards the aorta (Fig. 7(b)). This demonstrates, in concordance with the current pathological knowledge, the significantly higher sensibility of the aortic leaflets and ascending aorta.

A similar investigation conducted on the pulmonary trunk data set (Fig. 7(c)) showed that the discriminant regions are around pulmonary valve sinuses and at the right ventricular outflow tract. This clearly reflects the morphologic differences between the types of pulmonary trunks. Moreover this shows, that the relevant regions are much larger compared to the aortic valve, which also necessitates the use of larger numbers of features compared to the experiments with the aortic valve. Therefore the use of shape models and manifolds as subspace representations will be part of our future work, in order to find more representative local shape descriptors as features.

4. CONCLUSION

In this paper a framework for retrieval and decision support in the management of VHD was presented, which uses learning based discriminative distance functions. Various learning distance functions were evaluated including learning from equivalence constraints and the intrinsic Random forest distance. The generality and accuracy was shown by applying the method to different problems and large datasets from the domain of cardiology. Our method integrates the entire 3D geometrical expression of subject patient in an automatic classification framework offering a reliable and reproducible solution for diagnosis and decision support, which has the potential to advance the clinical management of valvular heart diseases.
Acknowledgements

The authors would like to thank the Medical University of South Carolina, the Ohio State University Medical Center, the Beth Israel Deaconess Medical Center, the Columbia University Medical Center, the Great Ormond Street Hospital for Children and the German Heart Center Munich for generously providing the images.

Part of this work has been done in the framework of the EU project Health-e-Child (IST 2004-027749) and the MEDICO Project in the THESEUS Program, which is funded by the German Federal Ministry of Economics and Technology under the grant number 01MQ07016. The responsibility for this publication lies with the authors.

REFERENCES


