Advanced Valve Modeling and Multi-Modal Fusion for Minimally Invasive Valve Procedures

Sasa Grbic

November 19, 2013

Technische Universität München
Department of Computer Science
Chair for Computer Aided Medical Procedures and Augmented Reality

campar.cs.tum.edu

TECHNISCHE UNIVERSITÄT MÜNCHEN
Lehrstuhl für Informatikanwendungen in der Medizin

Advanced Valve Modeling and Multi-Modal Fusion for Minimally Invasive Valve Procedures

Sasa Grbic

Vollständiger Abdruck der von der Fakultät für Informatik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzende: Univ.- Prof. Dr. Burkhard Rost
Prüfer der Dissertation:
1. Univ.- Prof. Dr. Nassir Navab
2. Prof. Terry M. Peters, Ph.D.
University of Western Ontario, Kanada

Die Dissertation wurde am 09.12.2013 bei der Technischen Universität München eingereicht. und durch die Fakultät für Informatik am 04.08.2014 angenommen.

Abstract

The treatment of valvular heart disease has shifted from open-heart surgery towards minimally invasive procedures. These procedures reduce procedural morbidity, mortality and treatment cost while accelerating patient recovery. However as there is no direct view or access to the affected anatomy, pre-operative planning in combination with procedural guidance is crucial for a successful outcome. Thus non-invasive imaging in combination with fast, precise and reproducible image analysis tools will become an essential part of clinical practice. This thesis will concentrate on the two main aspects of minimally invasive valve procedures: planning and guidance. We will present a novel dynamic patient-specific model of the complete valvular apparatus in combination with a robust machine learning framework to estimate the model parameters from computed tomography images. Advanced clinical measurements can be derived from our models and used for diagnosis, patient selection, implant selection and sizing. In addition we provide an advanced planning framework for the aortic valve implantation procedure (TAVI) where we combine a volumetric model estimation method with a virtual valve deployment framework. Further we introduce a robust model-based fusion framework to fuse high-quality pre-operative modalities with low-quality intra-operative images. In the first phase the pre-operative modality will be used to improve the model estimation accuracy in the low-quality intra-operative image. The second part involves developing novel multi-modal model based registration approaches to register pre-operative and intra-operative images. We employ robust machine learning techniques during the estimation. The methods are extensively validated on a large number of patients from multiple medical centers around the world. Both the planning and guidance frameworks have the potential to improve the treatment outcome while lowering procedural risks and treatment costs.

Zusammenfassung

Die Therapie von Herzklappenfehlern hat sich von offenen Herzoperationen auf minimal-invasive Behandlungen verlagert. Diese reduzieren die Morbidität, Mortalität und Behandlungskosten währenddessen die Genesung vom Patienten beschleunigt wird. Jedoch hat der behandelnde Arzt bei solchen Prozeduren keinen direkten Blick auf die betroffene Anatomie. Deswegen sind Planung und interventionelle Führung essentiell für einen erfolgreichen Eingriff. Nicht-invasive Bildgebungen in Kombination mit schellen, präzisen und reproduzierbaren Bildverarbeitungsalgorithmen werden essenzieller Bestandteil des zukünftigen klinischen Alltags werden. Diese Dissertation konzentriert sich deshalb auf die beiden Hauptaufgaben von minimal invasiven Herzklappenbehandlungen: Planung und interventionelle Führung. Wir präsentieren ein dynamisches, patientenspezifisches Model vom gesamten Herzklappenapparat in Kombination mit einem robusten System basierend auf maschinellem Lernen um das endgültige Model von Computertomographiebildern zu berechnen. Fortgeschrittene klinische Messungen können von unserem Modell abgeleitet werden und zur Diagnose, Patientenselektion, Implantatselektion und Messung herangezogen werden. Zusätzlich stellen wir ein Verfahren vor wobei die populäre minimal-invasive Aortenklappen-Intervention (TAVI) simuliert werden kann. Weiters präsentiern wir eine neue Methode vor für die Fusion von hoch-qualitativen präoperativen und niedrig-qualitativen intra-operativen volumetrischen Bildern. In der ersten Phase werden die prä-operativen Bilder dazu benutzt die niedrig qualitativen interventionellen Modelle zu verbessern. Danach stellen wir zwei neue Methoden fr̈ die Registrierung von Multi-Modalen Bilder vor. Hierbei verwenden wir robuste modelbasierte Methoden unter Zuhilfenahme von Methoden aus dem maschinellen Lernen. Alle vorgestellten Methoden sind ausführlich auf einer grossen Anzahl von Patienten aus mehreren klinischen Zentren quantitativ und qualitativ validiert worden. Beides die Planung und interventionelle Führung haben das Potential die jetzige Behandlung von Herzklappenfehlern zu verbessern wobei das Risiko und die Behandlungskosten reduziert werden.

Acknowledgments

First and foremost, I would like to thank my advisers Nassir Navab from TU Munich, Dorin Comaniciu and Razvan Ionasec from Siemens for their guidance and support during the last years of my dissertation. Nassir, thank you for accepting me at your prestigious group and providing me with your valuable advice. Dorin, thank you for your continuous support and your scientific insight and advice. I am especially thankful for your encouragement and support to visit clinical conferences and learn more about the real challenges clinicians face and how we can make a lasting impact with our research. I would also like to express my deepest gratitude to Razvan Ionasec, who guided me trough my first internship at SCR and introduced me to the exciting topic of heart valves. His passion and encouragement persuaded me to continue the work on this exciting topic.

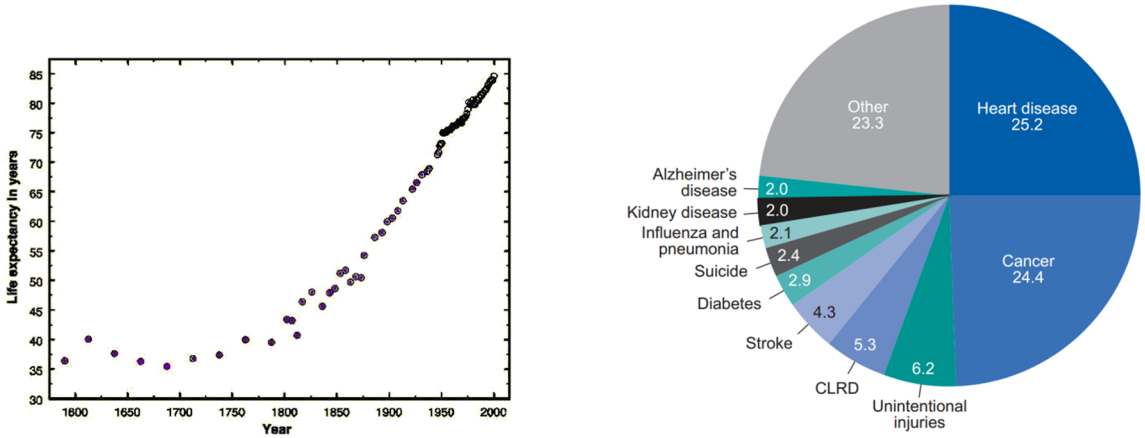
I had also the privileged to work with a number of great colleagues and students in the past years while residing both in Germany and the US. First I would like to thank my friends and colleagues Ingmar Voigt, Dime Vitanovski and Kristof Ralovich who were always friendly and supportive, acted as great team members along all these years. I would also like to thank Dirk Breitenreicher for his valuable comments during correction of this thesis. Many other colleagues at SCR have provided valuable discussions and input to this work, including: Tommaso Mansi, Bogdan Georgescu, Yefeng Zheng, Kevin Zhou, Yang Wang, Neil Birkbeck, Nathan Lay, Viorel Mihalef, Sharma Puneet. I am also thankful for the great students I was able to work with during the last years especially Dominik Neumann, Joshua Swee and Christian Gesell.

Finally I would like to thank my parents for their support and understanding during the long years of my education.

Contents

1. Introduction

During the last two centuries life expectancy has more than doubled [4, 144] (see Figure 1.1 left). The major contributors were tremendous advances in living standards, nutrition and medicine since the beginning of the industrial revolution in 1820. Especially advances in medicine with novel therapies including drugs, surgery and novel screening technology had a significant impact. Top leading diseases causing death changed frequently during this time. Currently cardiovascular and cancer are the most prevalent diseases accounting for almost 50% of all deaths in the developed world (see Figure 1.1). Cardiovascular diseases (Cardiovascular Disease (CVD)) are at the top with 500 000 deaths alone in the United States. In addition to higher mortality rates compared to cancer, treatment costs for CVD are significantly higher, 400 billion vs 250 billion.



(a) (b)

Figure 1.1.: (a) Life expectancy graph for the last four centuries in the developed world. (b) Relative comparison of major causes of death in the developed world. The statistics were retrieved from the Millenium Group Research [133].

1.1. Motivation

Valvular Heart Disease (Valvular Heart Disease (VHD)) is the most prevalent subgroup of CVD, affecting 2.5% of the global population and requires yearly over 100,000 surgeries in the United States alone and is a representative instance for the growing public health problem provoked by CVD. Heart valve operations are the most expensive procedures

with the highest mortality rate, an average cost of $141,120 and 4.9% in-hospital death rate [42, 14, 15].

Decisions in valvular disease management increasingly rely on non-invasive imaging techniques. The quality of acquired information, as well as the accessibility and cost effectiveness of each medical imaging modality has significantly improved over the past decades. Techniques like Computed Tomography (Computed Tomography (CT)), Transesophageal Echocardiography (Transesophageal Echocardiography (TEE)) and Cardiovascular Magnetic Resonance (CMR) imaging enable dynamic four dimensional scanning of the beating heart over the whole cardiac cycle. In addition modern C-arm systems are capable to acquire a three dimensional CT like modality within the operating room. Such volumetric time-resolved data encodes comprehensive structural and dynamic information, which however is seldom exploited in current clinical practice, due to its size and complexity as well as the lack of appropriate medical systems that utilize the extensive image information.

The progress in medical imaging is matched by important advances in surgical techniques, bioprosthetic valves, robotic surgery and percutaneous interventions, which have led to a twofold increase in the number of valve procedures performed in the United States since 1985 [91]. According to Millennium Research Group (MRG), the global authority on medical technology market intelligence, the United States heart valve market will grow strongly through 2016, almost entirely as a result of the introduction of devices for percutaneous valve therapies [133]. The market will reach a value of over $1.5 billion by 2016. Percutaneous or minimally invasive valve procedures tend to be less invasive, reduce procedural morbidity, mortality and the intervention cost while accelerating patient recovery. Powerful computer-aided tools for extensive non-invasive assessment, intervention planning, guidance and follow-up quantification are mandatory to continuously decrease the level of invasiveness, reduce the procedural risk and maximize effectiveness of the valve therapy.

1.2. Aims

The focus of this thesis is to develop efficient, precise and reproducible image analysis tools from non-invasive imaging modalities in order to enable future clinical applications in the area of minimally invasive valve therapies. The main emphasis is on pre-operative planning and intra-operative guidance. The following goals were pursued:

- Design an efficient mathematical model of the complete valular apparatus consisting of the aortic valve (Aortic Valve (AV)), mitral valve (Mitral Valve (MV)), tricuspid valve (Tricuspid Valve (TV)) and pulmonary valve (Pulmonary Valve (PV)) as well as the ascending aorta, which can capture complex anatomical, dynamical and pathological variations.

- Develop a approach for fast and robust patient-specific parameter estimation for the

complete valvular model (developed in previous Aim) from multi-phase CT images.

- Design a computational framework based on bio-mechanical models for advanced planning for the current most significant minimal invasive procedure: Transcatheter Aortic Valve implantation (TAVI).

- Develop a robust approach to register high-quality diagnostic data and low-quality intra-operative data for guidance of minimal invasive valve procedures (e.g. Transcatheter Aortic Valve Implantation - Transcatheter Aortic Valve Implantation (TAVI)).

1.3. Contributions

The major contributions of this thesis along with the corresponding publications:

- A hierarchical model of the complete valvular apparatus consisting of the aortic valve (AV), mitral valve (MV), tricuspid valve (TV) and pulmonary valve (PV) was proposed in [67] and [68] including the ascending aorta [70] which is able to capture complex anatomical and functional variation, including many valvular diseases. Consequently efficient machine learning techniques are incorporated to estimate the model parameters. In particular a new constrained Multi-linear Shape Model (cMSM), conditioned by anatomical measurements, is introduced to represent the complex spatio- temporal variation of the heart valves.

- A novel volumetric model for the aortic valve was proposed in [66] necessary for advanced planning for the TAVI procedure. We concentrated thus on a patient cohort with severely stenotic aortic valves. Based on our previous segmentation approach [67, 68] tissues within the aortic valve can be classified in one of the following categories: leaflet tissue, blood pool and calcification. In addition to intensity based features we incorporated novel geometrical features to capture the spatial context of each tissue in respect to the aortic valve anatomy.

- A computational modeling framework was developed to simulate the implant deployment procedure for TAVI [71]. From a cardiac CT a geometrical model is extracted and used to personalize the bio-mechanical computational model. In correspondence with a bio-mechanical model of the implanted device and corresponding boundary conditions (e.g. tissue strain/stiffness parameters, ventricular pressure) the deployment can be personalized. Thus different implant sizes and types can be easily tested prior to the actual procedure.

- To guide minimal invasive cardiac procedures two novel techniques were developed. Firstly an algorithm was presented to extract accurate valve models from low-quality contrasted intra-operative images by using additional pre-operative information [69]. Hereby a learning based approach is utilized where both models of the

same patient are estimated at once. Thus geometric models from low-quality interventional image can be improved by incorporating information from high-quality diagnostic images. The second guidance technique registers high-quality pre-operative images with low-quality non-contrasted intra-operative images [142, 65] by exploiting an efficient model based matching approach.

1.4. Outline of Thesis

The thesis consists of two major chapters, both focusing on minimally invasive procedures. The third chapter focuses on developing novel parametrization and estimation approaches of the complete valve apparatus for diagnostic quantification of valve anatomy and function and treatment planning. Further, a robust approach to guide minimally invasive procedures is explored whereby hi-quality pre-operative data is fused with low-quality intra-operative images in order to enable novel clinical guidance applications. A brief description of the subsequent chapters follows:

Chapter 2: Background

In this chapter we introduced the clinical background and the relevant imaging systems for cardiac diagnostic and interventional scanning. In the first part the main properties of the cardiac anatomy, with an emphasis on the valves, are explained. In the second part the importance of current diagnostic (CT, MRI, Ultrasound, IVUS, CTO) and interventional (C-arm CT, Ultrasound) imaging modalities are explored in respect to planning and treatment of valvular disease. In the last part we explore the current state-of-the art in medical image analysis related to our problems in model segmentation and multi-modal registration.

Chapter 3: Pre-operative Modeling and Quantification

This chapter describes in detail the proposed physiological model of the complete valvular apparatus and its parameter estimation. First a unified mathematical representation is introduced for all four heart valves and followed by the description of discriminative machine learning methods to estimate the model parameters. It is followed by improvements of modeling accuracy by presenting volumetric segmentation methods for the aortic valve for advanced intervention planning. In addition we utilize this approach in combination with a bio-mechanical modeling framework to simulate valve therapy. Lastly we present a model estimation approach from multiple modalities where we show that models in low-quality interventional images can be improved by utilizing information from high-quality pre-operative images.

Chapter 4: Intra-operative Guidance

In this chapter, a novel algorithmic framework is introduced to estimate the alignment between high-quality pre-operative images and low-quality interventional images. In con-

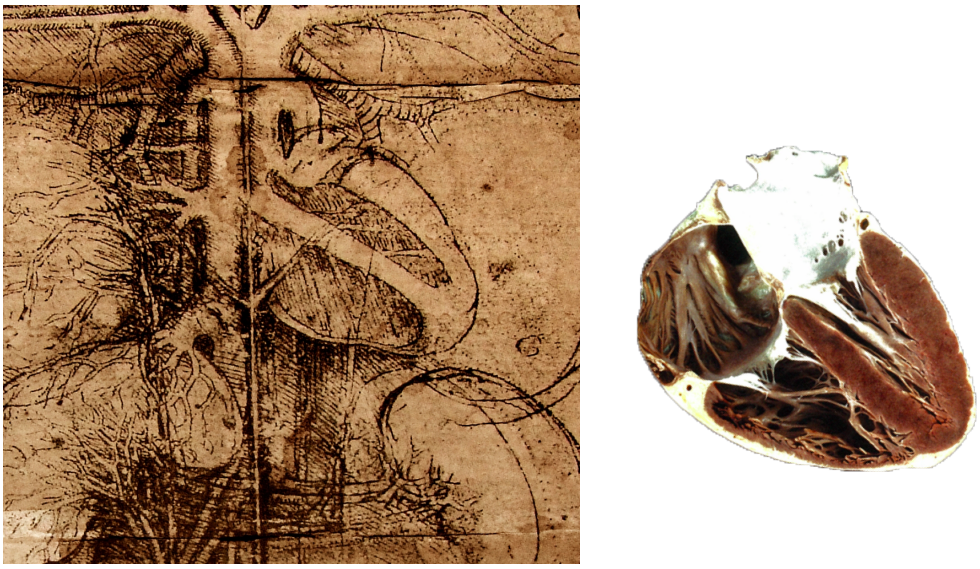
trast to intensity based registration methods we propose a model based registration approach which has the properties of retrieving a robust, fast and application specific image registration capable to cope with severe noise in the interventional setting.

Chapter 5: Conclusion

Finally we conclude in chapter 5 with a summary of the presented methods, their benefits and impact in the clinical environment, as well as a future outlook.

2. Background

Starting in the 16th century with Andreas Vasalisus the study of the heart started accelerating. In the subsequent centuries the field continued the pathway of descriptive anatomy and pathology, where many vividly depicted anatomy drawings surfaced such as Leonardo da Vinicis (see Figure 2.1). In the 19th and 20th century major advances have been made in the field of diagnoses and treatment of heart diseases. However the field cardiology was formed in the 20th century when first diagnosis instruments - blood pressure instrument, chest x-ray, the electrocardiogram (Electrocardiography (ECG)) were introduced.



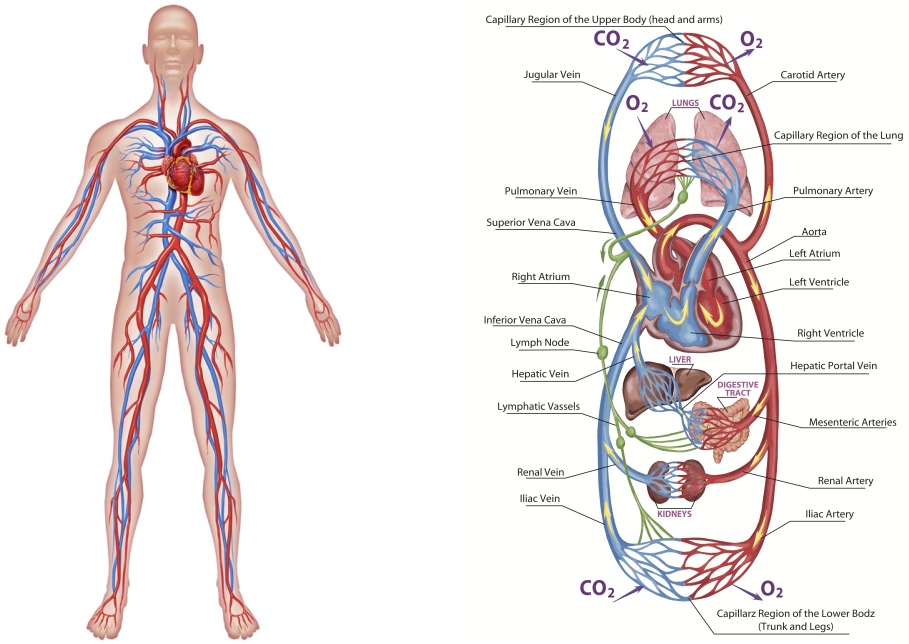
(a) (b)

Figure 2.1.: (a) Four-chamber view of the heart as illustrated by Leonardo da Vinci. Even the thinner wall size in the right ventricle was depicted correctly. (b) The four chamber view of the real heart from an ex-vivo specimen. Reproduced with permission of the European Association for Cardio-Thoracic Surgery. Multimedia Man Cardiothorac Surg doi:10.1510/mmcts.2006.002147.

In the subsequent sections the main anatomical and physiological properties of the cardiovascular system will be given. It will be followed by a description of the most common pathological disorders and their current treatment options. Finally the main imaging modalities for cardiac screening, intervention planning and guidance will be presented as well as the current advances in medical image analysis technology.

2.1. Cardiovascular system

The cardiovascular circulatory system is composed of the heart, blood vessels and blood. The heart acts as a pump and sends blood through elastic vessels called arteries, which branch into smaller ones and transport oxygenated blood through the body. The arteries divide finally into capillaries, with extremely thin walls, where oxygen, nutrients, minerals and other substances can pass through to surrounding cells and tissue. Waste substance, including carbon dioxide, flow from the tissue and cells into the blood for disposal. In a symmetrical analogy to the arteries the capillaries join to form vessels that eventually become veins. Thus they take blood back to the heart (see Figure 2.2 left).



(a) (b)

Figure 2.2.: (a) Diagram of the cardiovascular circulatory system. Vessels carrying oxygenated blood (usually arteries) are shown in red and those carrying deoxygenated blood (usually veins) are shown in blue. (b) Diagram showing the systemic and pulmonary circulation.

In the pulmonary circulation (see Figure 2.2 right) the right side of the heart pumps blood to the lung in order to oxygenate blood and then back to the left side of the heart. In the systemic circulation, the left side of the heart pumps oxygenated blood through the body tissues where it is depleted of oxygen and returned back to the right side of the heart.

2.2. The Human Heart

The heart is a powerful dynamic organ which pumps blood around the body's vast vessel network. It weights between 250 and 350 grams and is about the size of a fist. It consists of two main parts, the left and the right heart, which are intrinsically divided into two cavities, the atria and ventricles. Each side of the heart contains two valves regulating the blood flow both within the atria and ventricle but also the flow between ventricle and outflow vessels (see Figure 2.3).

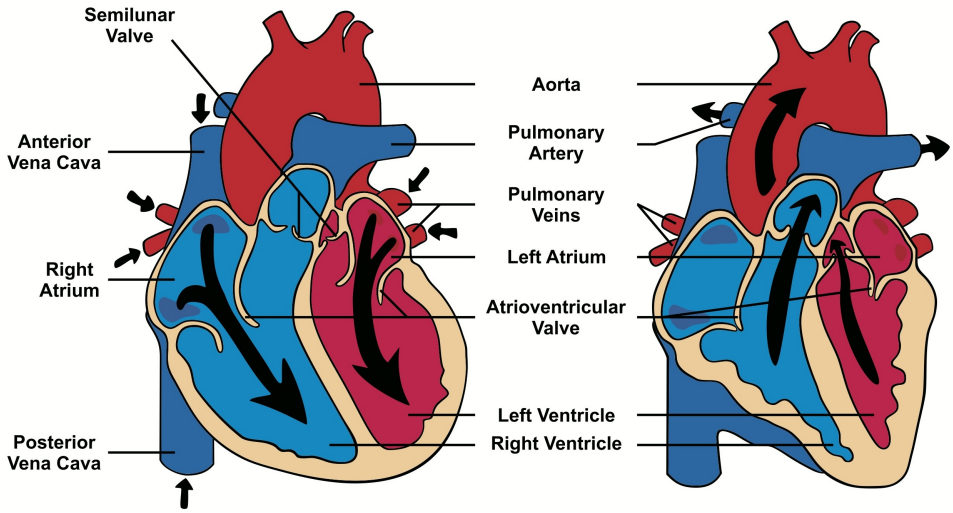


Figure 2.3.: Illustration of the heart during the main 2 cardiac phases: diastole (left) and systole(right).

The main power of the heart comes from the two lower chambers (ventricles) which have thicker muscular walls then the upper chambers (atrias). The ventricle contracting motion squeezes the blood into the arteries while the atrias act as passive reservoirs for blood flowing from the main veins. The average frequency of the heart is 60 beats per minute whereby each heartbeat can be divided into two main phases: systole and diastole. These two phases can be subdivided further, constituting four phases for the entire cardiac cycle: 1) relaxation (late diastole), 2) contraction of the atria (atrial systole), 3) contraction of the ventricles (ventricular systole) and 4) relaxation (early diastole).

During relaxation the muscular walls of the heart relax, the atrial chambers expand slightly as they fill with blood coming in under low pressure from the main veins. De-oxygenated blood from the body comes into the right atrium, while oxygenated blood enters the left atrium. Blood also flows from the atria into the ventricles, which are filled with 80% of their capacity.

The sinusoidal node is located in the located in the upper part of the right atrium. It sends electrical impulses which initialize the contraction phase. Impulses spread though the atrial walls and simulate their cardiac muscle to contract. This squeezes the blood

inside the atria through the atriventricular (mitral and tricuspid) valves into the ventricles which walls remain relaxed.

During the contraction phase (ventricular systole), the most active stage of the heartbeat, the thick cardiac ventricular muscle contracts, through the atrioventricular node. This causes a rise in the ventricular pressure, which opens both the aortic and pulmonary valves. Both valve are located at the outflow of the left and right ventricles.

In the relaxation phase the walls of the ventricles begin to relax, causing ventricular pressure to reduce. Due to the higher pressure in the outbound arteries the aortic and pulmonary valves close. This is preventing the back-flow of blood from the arteries into the ventricles. As the ventricular pressure reduces on the atrioventricular valves (mitral and pulmonary) the valves start to open. This reduces further the pressure in the atria, allowing blood to enter from the main veins (see Figure 2.4).

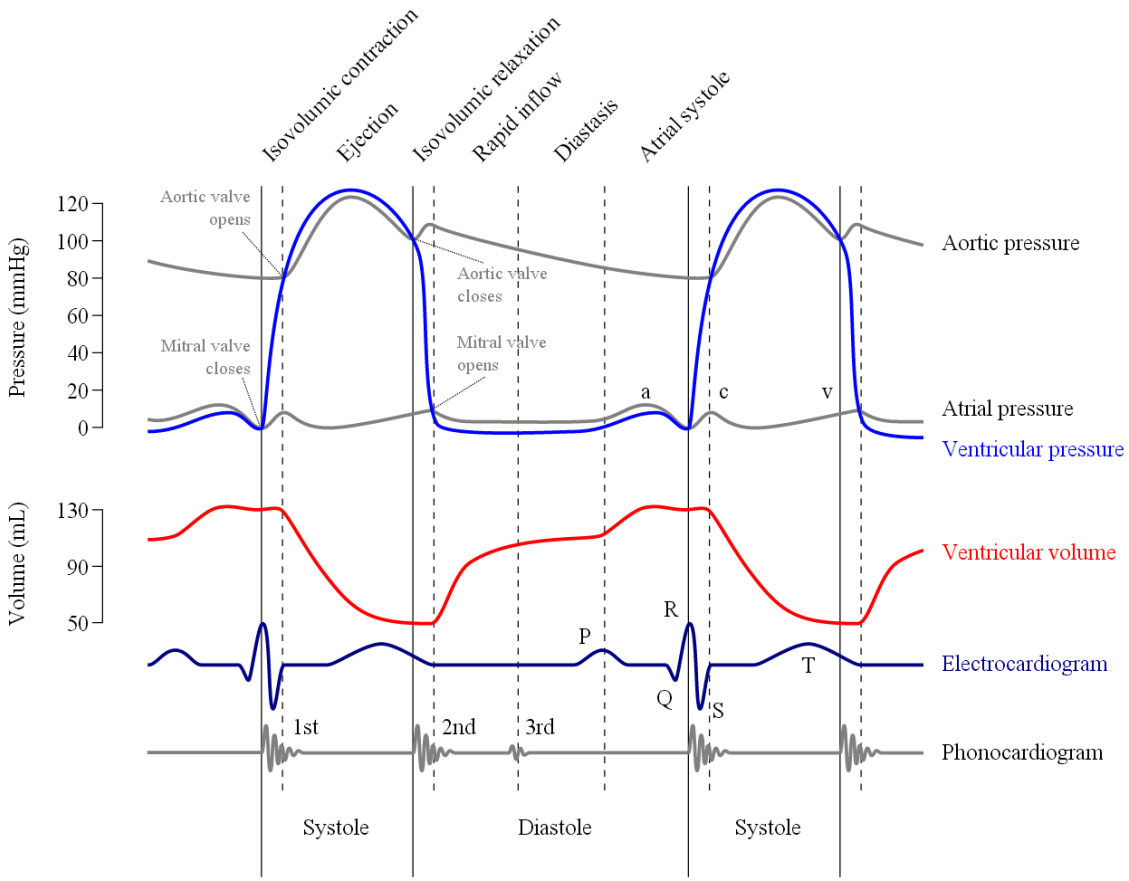


Figure 2.4.: Diagram of the left heart during the cardiac cycle (Wikipedia).

2.3. Physiology of the heart valves

Heart valves are structures within the human heart limiting the blood flow in one direction. The valves open and close passively depending on pressure differences on each side. They can be separated into two main groups: atrioventricular and semilunar valves (see Figure 2.5). The atrioventricular valves, consisting of the mitral and tricuspid valve, are thin structures composed of endocardium and connective tissue. They are located between the atria and ventricles. The Semilunar valves are flaps of endocardium and connective tissue reinforced by fibers to prevent the valves from turning inside out. As they are shaped as a half moon they are called semilunar. They are located between the left ventricle and the aorta (aortic valve) and the right ventricle and the pulmonary artery (pulmonary valve).

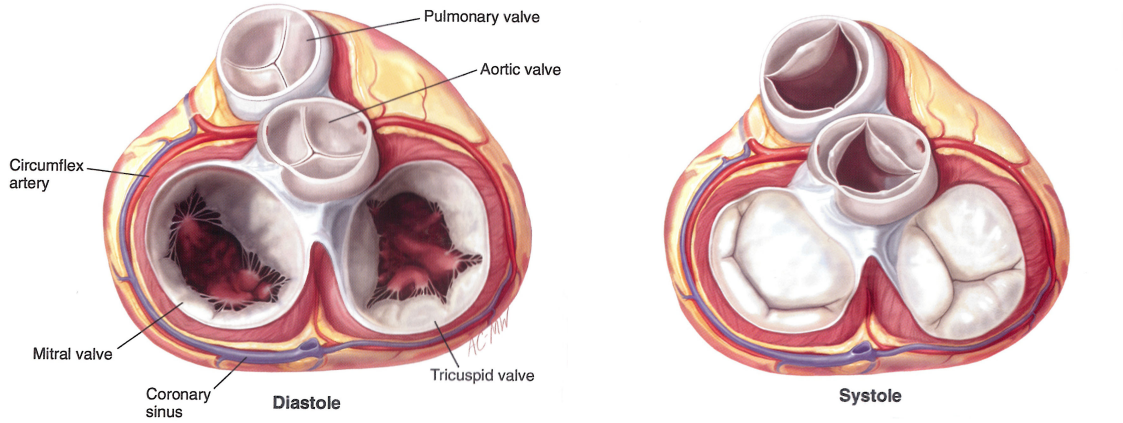


Figure 2.5.: The valvular apparatus during diastole (left) and systole (right) consisting of four heart valves: the aortic, mitral, tricuspid and pulmonary valve. Reproduced with permission from [26].

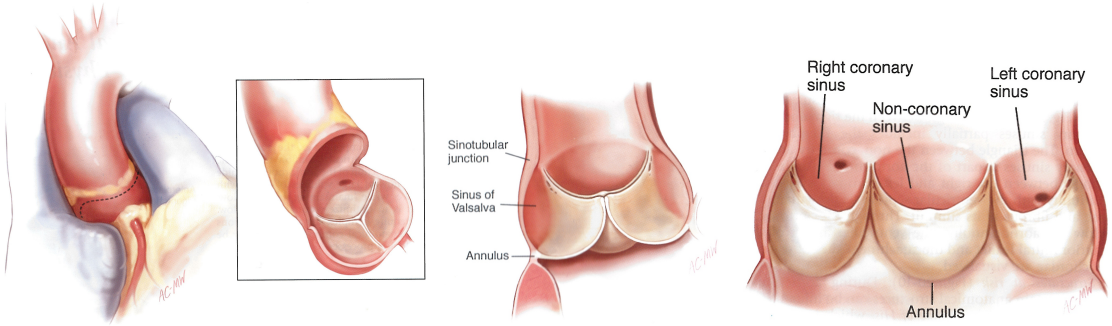
2.3.1. Aortic Valve

The central anatomical structures of the aortic valve consist of the aortic root and the three aortic valve leaflets, also referred as cusps. The aortic root is a tubular structure connecting the left ventricle outflow tract with the ascending aorta. Its function is to allow only for unidirectional blood flow from the left ventricle to the ascending aorta.

The lowest ring on the aortic root is called the annulus. The annulus incorporates the three hinge points where the three aortic valve leaflets are connected. The areas on the root which correspond to the leaflets are dilated and called valvular sinuses.

The aortic valve leaflets are attached to the aortic root within the valvular sinuses. The cusps form pockets with the lowest point called hinge. There are three hinge points, each cusp contains one.The attachments of each cusp ascend on the valular sinus where they

interlink at the level of the sinutubular junction forming the three commissures. Each cusp contains two commissures. The center of the free edge of the leaflet is refereed as the leaflet tip, consisting of a thicker fibrous tissue (see Figure 2.6).



(a) (b) (c)

Figure 2.6.: (a) Diagram of the aortic valve located between the left ventricle and the ascending aorta. (b) the aortic valve anatomy during diastole. (c) Unfolded view of the aortic valve anatomy emphasizing the three leaflets (cusps). Reproduced with permission of the authors from [26].

Underneath the sinotubular junction the two coronary ostia are located. They supply the left and right ventricle with oxygenated blood.

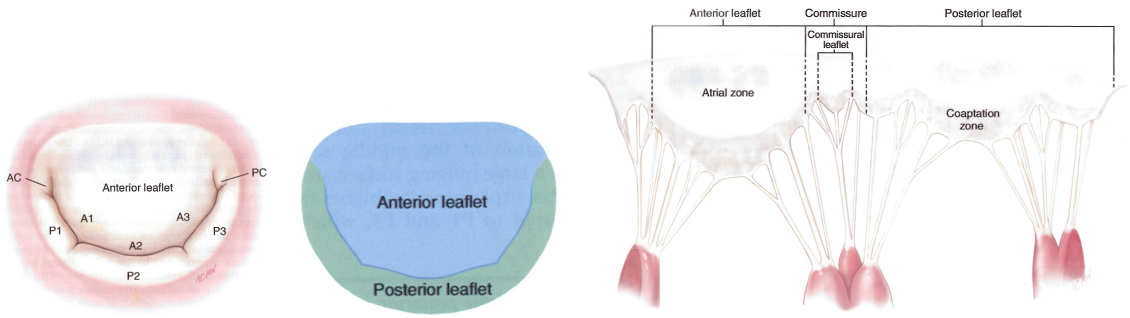
2.3.2. Mitral Valve

The central anatomical structures of the mitral valve consist of the two leaflets (anterior and posterior), the annulus, the papillary muscles and the chordae tendineae. The mitral valve is connecting the left atrium and the left ventricle and is supposed to prevent backflow of blood form the left ventricle to the left atrium during systole. The fibrous tissue on the outer perimeter of the anterior and posterior leaflet is called the mitral valve annulus.

The anterior leaflet is semi-circular shaped and separates the ventricular inflow and outflow tracts. In contrast to its right-sided counterpart, the tricuspid valve, it also forms part of the left ventricle outflow tract. The posterior is shaped as a crescent moon. Both leaflets can be divided into subparts called scallops. The anterior is divided into A1, A2, A3 while the posterior consist of the complementary scallops P1, P2, P3. In healthy patients both leaflets have similar areas. The anterior leaflet has almost twice the height compared to the posterior but only half of its annular length.

The commissures are cleft-like splits in the leaflet tissue and comprise the location where the posterior and anterior leaflets join. Below the commissures are the two papillary muscles, which arise from the left ventricular wall. Commissural chordae arise from each papillary muscle and extend in a fan-like array to insert into the free edge of both leaflets adjacent to the commissures (major commissures) or into two adjacent scallops of the posterior leaflet (minor commissures) [55]. As the chordae tendineae insert into the free-edge

and rough zone, they support the leaflets during systole to prevent prolapse (see Figure 2.7).



(a) (b) (c)

Figure 2.7.: Diagram of the mitral valve located between the left atrium and the left ventricle. (a,b) Top view of the mitral valve showing the anterior and posterior leaflet during systole. (c) Unfolded view of the mitral valve emphasizing the papillary muscles and chordae tendineae. Reproduced with permission of the authors from [26].

2.3.3. Pulmonary Valve

The pulmonary valve is one of the two semilunar valves located between the right ventricle and the pulmonary artery. The anatomical structure is similar to the aortic valve. It consists of a pulmonary root and three leaflets (cusps). Unlike the aortic valve which is continuously connected to the mitral valve the pulmonary and tricuspid valves are separated by infundibular muscle [55].

Analogous to the aortic valve the pulmonary valve opens due to increased ventricular pressure rise during systole. During diastole when the pressure in the right ventricle drops significantly, and thus the pressure gradients invert, the pulmonary valve closes.

2.3.4. Tricuspid Valve

The tricuspid valve is one of the ventricular valves. It is located between the right atrium and the right ventricle. The anatomical structure is composed of three aortic valve leaflets, the annulus and the subvalvular apparatus.

It is comprised of the annulus, leaflets, chordae tendineae, papillary muscles and commissures. In contrast to the mitral valve the tricuspid valve consists of three leaflets: posterior, anterior and septal leaflet. The anterior tricuspid leaflet is the largest and most mobile while the posterior is the smallest. The posterior leaflet is the least mobile because of its many chordal attachments.

2.4. Pathology of Valvular Heart disease

There are several types of heart disease. Many affect one valve and can be observed on all valves such as valvular stenosis and insufficiency [14, 184, 89].

Stenosis occurs when the valve opening is narrower than normal due to stiff or fused leaflets. This causes the heart to pump blood through the small opening and results often in heart failure. Valvular insufficiency also called regurgitation or incompetence manifests when the valve is not closing properly. Thus blood will leak backwards across the valve. In such cases the heart must work harder in order to compensate for the leaky valve [28, 187, 138, 25].

Congenital valve diseases affect mostly the aortic and pulmonary valve. Valves can be insufficient in size, have malformed leaflets which are often not attached properly to the annulus [82, 155]. Another congenital disease is the bicuspid aortic valve disease. Hereby instead of the normal three leaflets or cusps there are only two. In many cases the valve is unusually stiff causing stenosis or regurgitation [46].

Acquired valve diseases occur when valves which were once normal develop structural changes due to infections or other types of diseases. Hereby multiple valves can be affected. The most common forms are Rheumatic fever and Endocarditis [61, 60, 175, 38]. Rheumatic fever is caused by an untreated bacterial infection which usually occurs in children. Often symptoms are not observer for 20-40 years when the heart valves become inflamed, the leaflets stick together and become scarred, thickened and shortened. This usually leads to insufficiency. The introduction of antibiotics has dramatically reduced the number of this infection. Endocarditis occurs when bacteria enters the bloodstream and attack the heart valves causing scarring. This can lead to regurgitation. The bacteria enter the blood during surgery, dental procedures or severe infections.

Mitral valve prolapse is common, affecting 1-2% of the population. It causes the mitral valve leaflets to flop back into the left atrium during systole. Prolapse also causes the tissue of the valve to become abnormal and stretchy, causing insufficiency. However the insufficiency is mild in most cases and does not cause serious symptoms and usually does not require treatment.

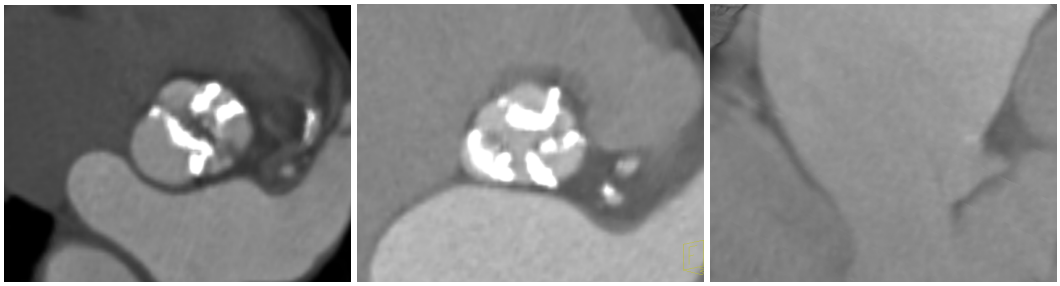
Other cause of valve disease can include coronary artery disease, cardiomyopathy (thickening of the left ventricle muscle), heart attack, hypertension and aortic aneurisms [7, 129, 12]. Many recent finding show that valvular disease is often correlated between valves [177, 188, 109]. Thus understanding and quantifying the whole valvular apparatus is a key prerequisite for precise and reproducible diagnosis of valvular disease.

2.4.1. Aortic Valve Stenosis and Insufficiency

In most cases aortic stenosis is caused by age-related progressive calcification of a normal aortic valve (see Figure 2.8). Other causes include congenital bicuspid aortic valve and acute rheumatic fever post-inflammatory.

The severity of aortic valve stenosis can be classified in four categories: 1) mild aortic

stenosis (<25 mmHg mean gradient and aortic valve area >1.5 mm^2), 2) moderate aortic stenosis (25-40 mmHg and 1.0-1.5 cm^2), 3) severe aortic stenosis (>40 mmHg and <1 cm^2) and 4) critical aortic stenosis (>70 mmHg and <0.6 cm^2)



(a) (b) (c)

Figure 2.8.: Examples of the aortic valve in CT with (a) bicuspid anatomy, (b) stenotic and heavy calcified aortic valve and (c) dilated aortic valve.

The primary cause of aortic valve regurgitation is aortic root dilation. In about 15% it is the cause of the innate bicuspid aortic valve and in another 15% cases are due to an inflammatory disease such as endocarditis. Based on the hemodynamic implications aortic insufficiency can be divided in two groups: acute and chronic insufficiency. Acute aortic insufficiency is considered a medical emergency with high mortality rates caused by endocarditis (see Figure 2.8).

2.4.2. Mitral Valve Stenosis and Insufficiency

Mitral valve stenosis is a disease characterized by a narrowing of the mitral valve orifice. In most case it is a consequence from a primary rheumatic heart disease. Other rather uncommon causes are calcification of the mitral valve leaflets and other forms of congenital diseases. The normal mitral valve orifice is between 4-6 cm^2. Mitral valve stenosis is usually classified in three categories: 1) mild (mean gradient <5 mmHg and opening area >1.5 cm^2 during diastole), 2) moderate (mean gradient between 5-10 mmHg and opening area of 1-1.5 cm^2 during diastole) and severe (mean gradient of >10 mmHg and opening area < 1 cm^2 during diastole).

Mitral valve insufficiency is the most common form of valvular heart disease. Based on the severity of the insufficiency it is usually separated in four groups classified based on the regurgitation fraction. The regurgitation fraction is defined as the percentage of the left ventricular stroke volume that regurgitates into the left atrium. The four groups of mitral valve insufficiency are are: 1) mild (regurgitation fraction <20%), 2) moderate (regurgitation fraction 20-40%), 3) moderate to severe (regurgitation fraction 40-60%) and 4) severe (regurgitation fraction >60%).

2.4.3. Pulmonary Valve Disease

Pulmonary valve disease is primarily associated with congenital lesion often including the right ventricular infundibulum. It is rarely associated with an infectious process such as rheumatic fever or bacterial endocarditic. It can be pathologically affected in carcinoid heart disease with resulting insufficiency or stenosis.

2.4.4. Tricuspid Valve disease

The tricuspid valve disease is often in association with or secondary to mitral or aortic valve disease or left ventricular (LV) disease. In contrast to the aortic or mitral valve disease tricuspid valve disease is rarely treated and often ignored.

2.5. Treatment of Valve Disease

Depending on the severity of the valvular disease and the physical condition of the patient there are three treatment options: 1) medication, 2) heart valve surgery and 3) minimally-invasive or percutaneous procedures.

In some cases for patients with mild valvular heart disease medications can be an option. Commonly prescribed medications for valvular heart disease include medications to open blood vessels (vasodilators), medications to lower cholesterol (statins), medications that reduce water retention (diuretics) and blood-thinning medications (anticoagulants). In addition to patients with mild valular heart disease patients with extremely high risk-scores, deemed inoperable, are usually treated with medications.

(a) (b) (c)

Figure 2.9.: (a) Carpentier-Edwards Bioprosthesis for the aortic valve, (b) The Carpentier-Edwards Classic annuloplasty ring and (c) Edwards ETlogix annuloplasty ring.

Most patients with valvular heart disease undergo surgery. Since the 1950 and the invention of the heart-lung machine the field of cardiac surgery was established. Repair and replacement procedures are regularly performed on low and medium risk patients measured using the standard risk assessment scores STS and EUROSCORE 2 [156, 80, 105, 157].

In the recent years a new way of treatment options has emerged. These procedures are called minimally invasive or percutaneous interventions. With novel and catheter devices with smaller footprint these procedures have received an exponential growth (see Figure 2.12). In comparison to standard surgical procedures these interventions can be utilized in high-risk and inoperable patients.

Despite the tremendous advances in treatment options costs for valvular interventions have been consistently high with an average cost of $ 141 200 per intervention and 4.9% in-hospital death rate. Minimal invasive procedure have the potential to lower the interventional cost constituting now more than 40% of all valvular interventions while the market is still expected to grow till 2018 [133]. However as there is no direct view or access to the affected anatomy pre-operative planning in combination with procedural guidance is crucial for a successful outcome. Thus non-invasive imaging in combination with fast, precise and reproducible image analysis tools will become an essential part of clinical practice to assure a successful outcome.

2.5.1. Heart Valve Surgery

All valve surgeries are performed under general anesthesia. The surgeon makes a large surgical cut in the breastbone to reach the heart and aorta. Most people are connected to a heart-lung bypass machine. The heart is stopped during the intervention while the heart-lung bypass machine replaces the heart function.

2.5.2. Valve Repair

Repair procedure are not widespread for every valve. The mitral valve repair procedures are the most common. Patients are usually suffering from prolapse and severe regurgitation caused either by dilating mitral valve annulus or defected chordae tendineae. In most cases a ring prosthesis (annuloplasic) ring, such as the Carpentier-Edwards Classic annuloplasty ring, is applied to remodel the mitral valve annulus. The available annuloplasty rings are rigid, flexible, complete, partial, and semi-rigid/flexible. Several objectives exist for annuloplasty, namely remodeling of the length and shape of the dilated annulus, prevention of dilatation of the annulus, and support for the potentially fragile area after partial-leaflet resection. Annuloplasty rings may have the potential for maintaining the anatomical and physiological characteristics of the mitral annulus. In most procedures the annuloplasty is combined with resections on the leaflets. In addition artificial chords, such as Neochordae by Valtech Cardio or Yehuda, Israel, can be used to re-attached the leaflets to the papillary muscles. Compared to mitral valve replacement the complexity of the procedure is very high and requires an experienced cardiac surgeon. Proper planning for this procedure is crucial and image analysis with combined computational modeling can help to improve procedural outcome, especially in lower volume clinical centers.

Aortic valve repair is seldom. In most cases it is associated with congenital disease such as the Ross operation where the dysfunctional aortic valve is replaced with the functional

pulmonary valve. However due to the high complexity of the procedure it is executed solely in highly specialized centers. Other procedures include commissuroplasty, where the valvular orifice can be narrowed in order to reduce aortic valve regurgitation.

Reconstructive surgical procedures for the right side valves are rare. For the tricuspid valve surgical procedures include suturing, repair with flexible rings and complete remodeling rings. However the evaluation and treatment selection varies wildly between clinical centers.

Valve Replacement

During the valve replacement the deficient valve is replaced by an artificial valve implant. There are two main implant groups: mechanical valves and tissue valves. Mechanical valves are more robust and last longer compared to tissue valves. However they contain increased risk of blood clots and thus require the intake of anticoagulants for the rest of a patients live. The tissue valves are usually made from animal valve tissue or animal pericardial tissue. In some cases homografts are used where the tissue is provided by human donors.

New bioprostetic aortic valves such as the Carpentier-Edwards bioprosthesis offers improvements such as reduced calcification risk, an improved fixation technique which minimizes alterations in the collagen waveform and increases leaflet compliance and minimized tissue stress (see Figure 2.12).

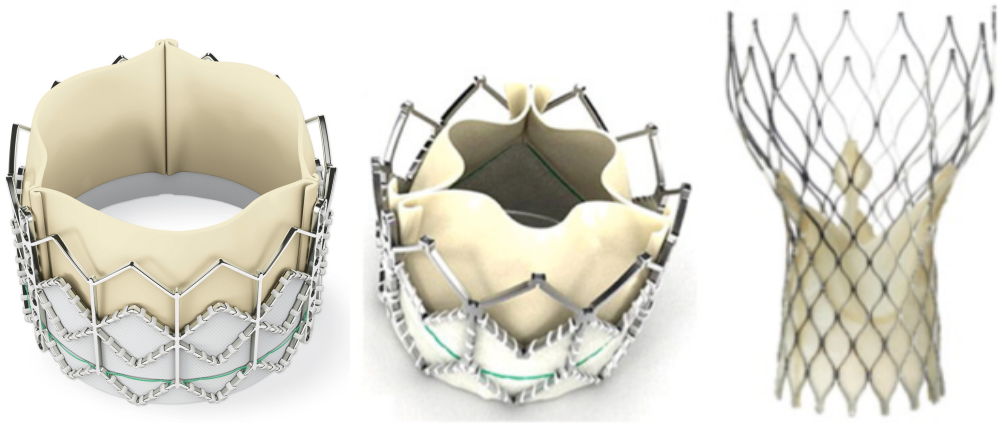
Minimally Invasive Procedures

In recent years minimally invasive or percutaneus procedures have replaced surgical repair and replacement [112, 99, 169, 93, 77, 177, 47]. These procedures offer the potential to reduce procedural morbidity, mortality, and costs of valve treatment. Many patients with systematic valular diseases were deemed as inoperable based on their risk assessment (STS score, Euroscore 2, etc.). Thus they could not undergo regular valve repair and replacement. Their only treatment option was blood thinning medication. Many of these patients can however undergo minimal invasive procedures. They are establishing themselves as a viable alternative to standard open-heart surgery.

During a minimally invasive procedure an implant is delivered through a catheter inserted through a small incision point in the vascular system (e.g. femoral artery). The procedures are executed in specialized operating rooms called Hybrid Operating Rooms or Cath-labs equipped with advanced imaging technology such as Fluoroscopy and Transesophageal Echocardiography (TEE). As there is no direct access and view to the affected anatomy these imaging modalities are installed to guide the procedure.

One of the most prevalent minimally invasive procedures is the transcatheter valve implantation (TAVI) where a replacement valve is delivered via a catheter using one of several access methods: transfemoral, transapical, subclavian and direct aortic.

The first TAVI procedure was performed in 2003 by Alain Cribier and within one decade the number of procedures increased to more than 50,000 implants done within 40 countries



(a) (b) (c)

Figure 2.10.: (a) The Edwards SAPIEN balloon-expandable valve. (b) The new SAPIEN XT. (c) The Medtronic CoreValve.

[34]. Many manufacturers such as Edwards Lifesciences (Irvine, California), Medtronic (Minneapolis, Minnesota), St. Jude Medical (St. Paul, Minnesota), Boston Scientific Inc. (Natick, Massachusetts), Direct Flow (Santa Rosa, California), JenaValve (Munich, Germany) etc. have released multiple devices with European CE marks and some have gained also the FDA approval. The two most popular implants are the SAPIEN and COREVALVE. The Edwards SAPIEN balloon-expandable valve incorporates a stainless steel frame, bovine pericardial leaflets and a fabric sealing cuff. The new version of the SAPIEN valve (SAPIEN XT) uses a cobalt chromium alloy frame and is compatible with lower profile delivery catheters. The Medtronic CoreValve incorporates a self-expandable frame, porcine pericardial leaflets, and a pericardial seal.

Other popular minimal invasive valve procedures include the mitral valve MitraClip system (Abbott Vascular, Santa Clara, CA, USA) where a selective group of patients can be treated. The prolapsed mitral valve, often caused by a broken chordate string, is repaired using a clip stitching the anterior and posterior leaflet (see Figure 2.11).

2.6. Imaging Modalities

2.6.1. Computed Tomography

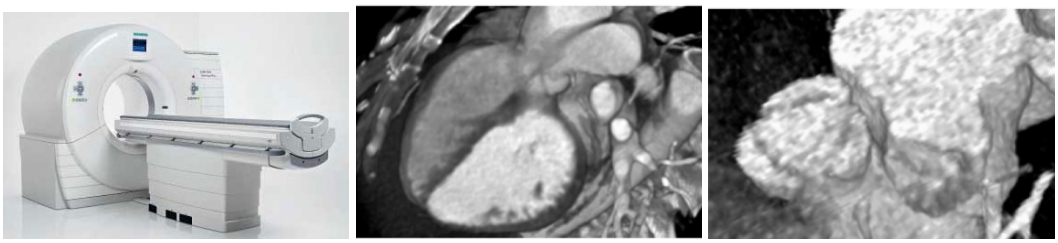
Computed tomography or a CT scan (also known as CAT scan) is an imaging technique that uses a high amount of ionizing radiation in conjunction with an reconstruction algorithms to recover a high-quality 3D diagnostic image of the patients internal anatomy. It can be utilized to fully evaluate both cardiac structure and function. Advances in the recent decades have allowed for a more complete evaluation of both stationary structures, such as the thoracic aorta, as well as rapidly moving structures such as the heart valves. Usually

(a) (b) (c) (d) (e)

Figure 2.11.: (a) Lotus valve (Boston Scientific), (b) Direct Flow valve (DirectFlow), (c) Portico valve (St. Jude Medical), (d) Engager valve (Medtronic Inc., Minneapolis Minnesota), (e) JenaClip valve (JenaValve).

cardiac acquisitions were combined with an Iodine based contrast agent in combination with ECG gating to ameliorate the effect of blur caused by cardiac motion. However with the new scanner generation such as the Siemens SOMATOM Definition Flash CT, Siemens Healthcare, Forchheim, Germany can acquire a temporal resolution of 75 milliseconds and require 0.6 seconds for a complete thorax scan at a radiation dose of 1 millisievert (mSv). This is done without ECG gating, increasing the efficiency, patient convenience and lowering the cost per exam. These advances in spatial and temporal resolution have helped in the evaluation of cardiac structures including coronary veins, pulmonary veins, atria, ventricles, aorta and the valvular apparatus (see Figure 2.12).

Computed tomography has gained a lot of interest in the advent of minimally invasive procedures [168]. For TAVI CT is currently the standard planning modality. Both implant type and size will be selected based on the patient anatomy measured from CT. In contrast to other modalities such as ultrasound patient specific parameters are more accurate [143].



(a) (b) (c)

Figure 2.12.: (a) CT - SOMATOM Definition Flash, Siemens Healthcare, Forchheim, Germany. (b) Complete heart scan with contrast on the left side. (c) Volumetric reconstruction of the aortic valve clipped at the valvular sinuses level.

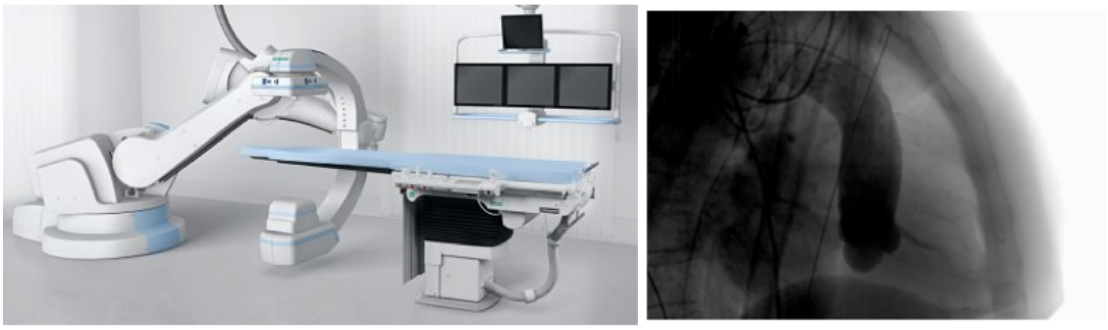
2.6.2. Angiography and Radiography

Radiography is an imaging technique where electromagnetic ionizing radiation is used to visualize the human body. In the context of cardiac imaging a chest X-ray can give information about the size and the configuration of the heart and the great vessels and also on the lung fields and vessels. It is a routine cardiac investigation procedure with very low radiation exposure.

Fluoroscopy, a term invented by Thomas Edison, provides a moving projection radiographs with lower ionizing radiation and thus quality. In combination of contrast agent it allows to observe soft tissue movement or to guide interventional procedures. Angiography involves the usage of fluoroscopy to visualize the cardiovascular system. An iodine-based contrast with high density is utilized to view the vessels under X-ray in order to find stenosis, aneurisms and valvular leakages.

Modern C-arm systems such as the Siemens zeego, Siemens Healthcare, Forchheim, Germany, offer the ability to acquire a CT like 3D acquisition called DynaCT (or 3D C-arm CT) in the operating room. During a 5 second and 200deg sweep 512 projections are acquired and a 3D volume can be reconstructed similar as computed tomography. Several acquisition protocols can be utilized to visualize specific parts of the vascular system. For TAVI a specialized protocol was designed to visualize both the aortic valve structure and the ascending aorta where contrast agent is injected at the level of the aortic valve annulus in combination with rapid pacing [94] during the C-arm sweep. Combined with advanced segmentation algorithms this image can be used to guide TAVI procedures (see Figure 2.13).

The largest drawbacks of DynaCT are the ionizing radiation which is usually larger than CT. In addition the contrast agent can be responsible for kidney failures and the usage of rapid pacing is putting additional stress on the patient during the intervention.



(a) (b)

Figure 2.13.: a) C-arm X-ray - Artis zee Ceiling-mounted system, Siemens Healthcare, Forchheim, Germany. (b) Fluoroscopic image of the heart with contrast concentrated in the aorta.

2.6.3. Echocardiography

Echocardiography is a routine exam used for quantification, diagnosis, management and follow-up examination of patients suspect of cardiovascular disease. An ultrasound machine transmits sound pulses into a patient's body using an array or matrix mounted on a transducer probe. These high-frequency sound pulses, usually ranging from 1 to 5 MHz travel through the patient's body, eventually hitting boundaries between tissues and reflecting back to the transducer probe. The tissue boundaries may be between soft tissue (connecting or surrounding organs) and fluids, such as blood, or between soft tissue and bone. Using the reflected signal and sophisticated reconstruction algorithms either 2D or 3D images can be extracted. Based on the type of probe and acquisition position used there are 2 main scan types of echocardiography for cardiology application: Transthoracic echocardiogram (Transthoracic Echocardiography (TTE)) and Transesophageal echocardiogram (TEE).

Transthoracic echocardiogram (TTE), or cardiac ultrasound is known as standard echocardiogram. In this case, the echocardiography transducer (or probe) is placed on the chest wall (or thorax) of the subject, and images are taken through the chest wall. This is a non-invasive, highly accurate and quick assessment of the overall health of the heart. Smaller structures such as the heart valve can also be assessed but the imaging quality is limited (see Figure 2.14).

Transesophageal echocardiogram (TEE) is an alternative way to perform an echocardiogram. A specialized probe containing an ultrasound transducer at its tip is passed into the patient's esophagus. This allows image and Doppler evaluation from a location directly behind the heart. This is known as a transesophageal echocardiogram, or TOE (TEE in the United States). Transesophageal echocardiograms are most often utilized when transthoracic images are suboptimal and when a more clear and precise image is needed for assessment. This test is performed in the presence of a cardiologist, registered nurse or ultrasound technician. Conscious sedation and/or localized numbing medication, may be used in order to make the patient more comfortable during the procedure (see Figure 2.14).

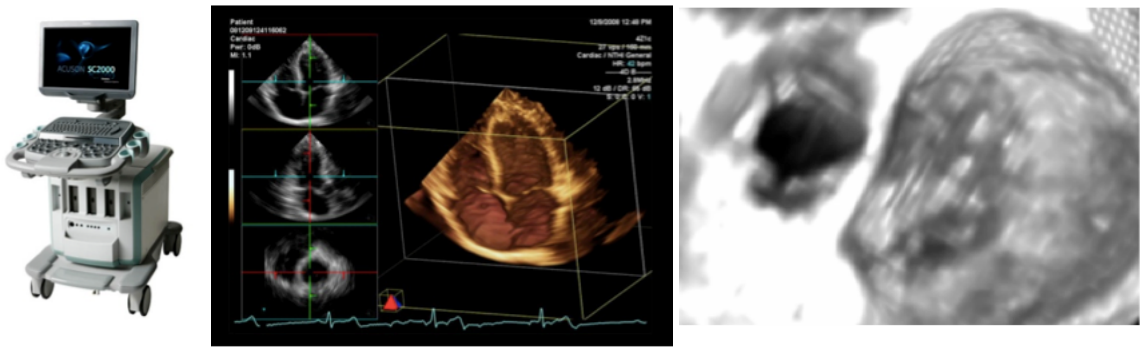
Compared to other diagnostic modalities ultrasound is almost risk-free and cost efficient diagnostic modality. Thus, it is the most common diagnostic test in cardiology.

The drawbacks of echocardiography is its low signal to noise ration in combination with a relative small field of view.

2.6.4. Magnetic Resonance Imaging

Magnetic resonance imaging (Magnetic Resonance Imaging (MRI)) is an imaging technique where a powerful magnetic field, either 1.5 T or 3T, radio-frequency pulse is emitted and absorbed by patient's hydrogen nuclei. This process can be measured and transformed into images (see Figure 2.15).

Static and dynamic images can be generated to assess many diseases and conditions, including coronary heart disease, heart failure, heart valve problems and congenital heart

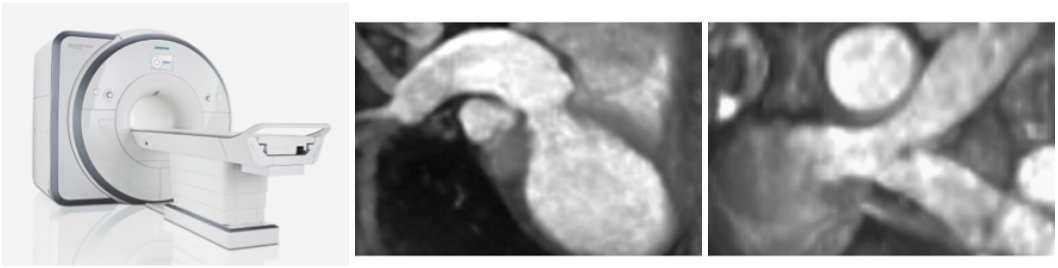


(a) (b) (c)

Figure 2.14.: (a) Siemens SC2000 ultrasound machine, (b) TTE four chambers image of the heart, (c) TEE volumetric reconstruction of the mitral and aortic valves.

defects. The advantages of MRI is the high contrast between soft tissues whereby no contrast agent is necessary. Contrast agent, such as gadolinium, might be injected during cardiac MRI, which increases the contrast in the heart and blood vessels on the MRI pictures. This contrast agent often is used for people who are allergic to the dyes used in CT scanning.

The largest advantage compared to CT and X-ray is that MRI does not emit any ionizing radiation and is thus almost risk free. The main disadvantage is the acquisition time which is usually 4-8 times higher than CT. In addition the acquisition protocol is complex compared to CT and requires an experienced user.



(a) (b) (c)

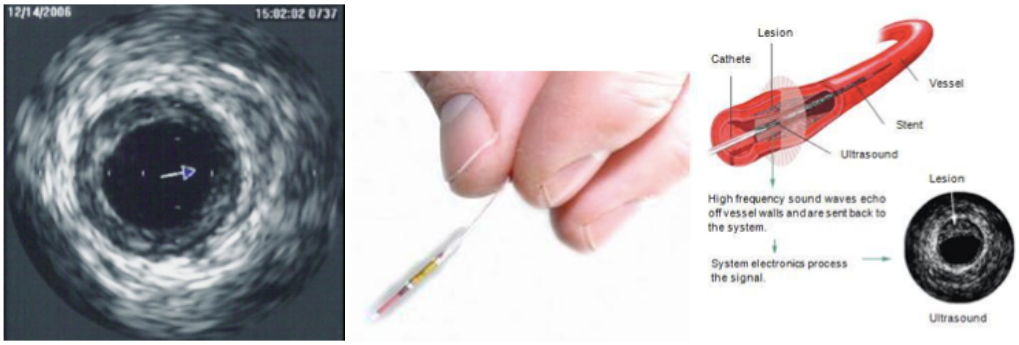
Figure 2.15.: (a) MRI - Siemens Somatom Spectra, Siemens Healthcare, Erlangen, Germany. (b) Heart image including the left ventricle and aorta. (c) Volumetric reconstruction of the right ventricular outflow tract and pulmonary arteries.

2.6.5. Intravascular ultrasound

Intravascular ultrasound (Intravascular ultrasound (IVUS)) is an imaging technique where a miniaturized ultrasound probe is attached to the end of a catheter (see Figure 2.16). With a diameter of 9F (3mm) the catheter can be inserted in small blood vessels in order to assess plaque and stenosis. The main application is coronary artery disease where the IVUS catheter is used to determine the amount of plaque built up at any particular point in the epicardial coronary artery wall (see Figure 2.16).

Compared with angiography it has several advantages. The most prominent is the visualization of plaque. Angiography can underestimate the infarction risk as increased plaque deposits will not be visible due to the simultaneously vessel diameter increase. In addition IVUS can be used to visualize multiple overlapping arterial segments. The catheter can also be used to assess the effects of stenosis treatments such as angioplasty and the results of medical therapy over time.

The major disadvantages of IVUS are the risk associated with catheterization and the increased examination time and its associated costs.



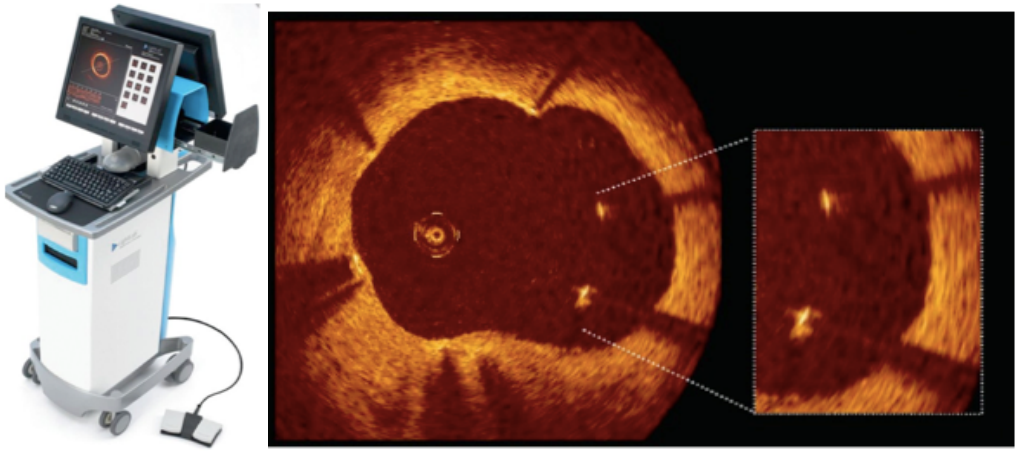
(a) (b) (c)

Figure 2.16.: (a) acquired image using IVUS of the coronary arteries, b) IVUS catheter, c) diagramm of IVUS working principle.

2.6.6. Optical coherence tomography

Recently, Optical Coherence Tomography (Optical Coherence Tomography (OCT)) has emerged as one of the most promising imaging modalities for cardiac diagnosis, especially for the coronary vessels due to its excellent image quality and small catheter dimension. OCT uses reflected light to create cross sectional images of the vessel (see Figure 2.17).

Compared with IVUS it offers an axial resolution 100x higher, 15 μm compared with 150 μm. In addition the probe size is half the size of the IVUS. During the acquisition the blood must be removed as it would cause multiple light scattering and attenuation. However new acquisition protocols have also emerged with non-occlusive techniques. Complications are seldom, though can include ischemia, arrhythmias and thrombus formation.



(a) (b)

Figure 2.17.: (a) Console OCT LightLab M3CV system (b)OCT appearance of a drug eluding stent 9 months after implantation. The amplified image shows two struts not opposed to the vessel wall.

2.7. Overview of Medical Image Analysis

Medical image analysis is an interdisciplinary field at the intersection of computer science, mathematics, engineering and medicine. The most general objective is to develop mathematical and computational algorithms to solve problems pertaining to medical images and their use for biomedical research and clinical care. One of the main goals is to extract clinically meaningful information from medical images. The field can be separated into several broad categories: segmentation, registration, image reconstruction, physiological modeling etc. In this thesis we focus on the topic of image segmentation and registration in the context of minimally invasive valve procedures. In the next three sections a short review of segmentation, computational modeling and registration algorithms will be provided.

2.7.1. Image Segmentation

Segmentation is the process of partitioning an image into different regions or segments. In medical imaging, these segments often correspond to different anatomies such as organs. Based on the method employed a classification in three broad categories can be made: atlas based approaches, model based approaches and interactive segmentation methods.

Atlas based approaches rely on a small set of training data where the anatomy of interest is manually segmented by an expert. During test time the unseen image will be registered to an annotated reference image and the annotation will be propagated from the reference frame to the new image. In the area of cardiology, Rikxoort et al. proposed an atlas-based

segmentation approach for the heart chambers and the aorta [186]. Hereby multiple atlases are registered to a target image. The final segmentation of the target is obtained by propagating atlas labels to the target image. The multiple propagated labels are combined using spatially varying decision fusion weights. Sabuncu et al. presented a label fusion approach within the area of atlas based segmentation [159], in which a weighted voting is formulated in terms of minimizing the total expectation of labeling error, and in which pairwise dependency between atlases is explicitly modeled as the joint probability of two atlases making a segmentation error within a particular region. This probability is approximated using intensity similarity between a pair of atlases and the target image in a local neighborhood.

Model based segmentation methods use a parametrization of the anatomy of interest. This can be achieved by using bounding boxes, landmarks or surface boundary points. Two of the most common shape-based techniques are Active Shape Models [33] and Active Appearance Models [32]. The active shape models try to fit a shape to an example of the object in a new image. Thereby an iterative approach is pursued whereby an alternating strategy is utilized: 1) find an update for each point in a small neighborhood in the current shape based on the image 2) update the model parameters to best match to these new found positions. Several approaches of this technique have been successfully applied in the area of medical imaging [40, 158, 185].

Machine learning techniques have been successfully applied to medical imaging problems since more than a decade [124]. However in recent years the usage of such techniques has proliferated, especially in the field of object detection and segmentation. With large medical databases become available, discriminative learning methods with semantic constraints have proven to be the solution of choice to solve estimation problems in high dimensional spaces. Challenging detection and segmentation tasks in images with low signal-noise ratio containing artifacts and signal dropout (such as TEE) became feasible. Georgescu et al. [58] introduced a boosting based approach to segment the left ventricle from challenging 2D ultrasound images. The usage of Marginal Space Learning was introduced by Zheng [205] as an efficient way of learning high dimensional problems by operating in spaces of increasing dimensionality. Thus a segmentation of multiple organs can be achieved within seconds with a high accuracy [205]. Criminisi et al. [35, 103, 59] proposed several approaches to locate multiple organs in CT images by using Random Forrests. Pauly et al. [148] utilized a random ferns based regressor to predict bounding boxes of organs from MRI images. Lindner et al. proposed a random forest approach with a statistical shape model to segment the Proximal Femur from X-ray images [114].

Several methods have been proposed which advance the model parametrization in comparison to the Active Shape Models [33]. Zhu et al. introduced a new shape representation based on multi-linear models [209, 210]. In this framework the temporal and geometrical shape information would be modeled independently in contrast to standard statistical shape models. The method was demonstrated on left ventricle segmentation from both MRI and [209] echocardiography [210]. Several techniques employ non-linear manifold

learning to describe both the temporal and spatial variation of shapes [201, 127]. Yang et al. derived the temporal components and dynamics of the heart model explicitly from patient scans by Motion Manifold Learning [201].

In the context of cardiovascular diseases many groups in medical image analysis have been focusing on the construction of patient-specific anatomical models from well established diagnostic modalities (e.g. CT and MRI) to aid disease analysis and treatment planning [68, 97, 211, 9, 3]. Zheng et al. proposed an efficient method based on machine learning techniques to extract patient-specific models of the heart chambers from CT. Within three seconds detailed geometrical models of the left ventricle, left atrium, right ventricle and right atria can be extracted [205]. Fritz et al. proposed a framework to segment the left and right cardiac ventricle using a combined bi-temporal statistical model [54]. Berg et al. [192] proposed a segmentation method of a geometric cardiac model including the four cardiac chambers and the trunks of the connected vasculature, as well as the coronary arteries and a set of cardiac landmarks from cardiac CT. A mean model based on an atlas is used with a consecutive refinement step where the model is personalized.

In the context of valvular disease management, Grbic et al. [68], Ionasec et al. [88] and Waechter et al. [195] proposed the modeling of the aortic valve from cardiac CT. Models of the mitral valve from MR have been proposed by Linte et al. [115]. Burlina et al. [23] proposed an interactive algorithm based on thin-tissue detection and level-set deformable models to identify the mitral valve and the left ventricle endocardium in 3D TEE images for an automated and efficient mitral valve assessment. Schneider et al. proposed a complex pipeline to automatically delineate the MV from 3D+t TEE images. The method relied on mitral annulus detection and tracking [166], leaflet segmentation of the open valve [164] and leaflet tracking using a deformable model that handled contacts and chordae stresses [165]. Temporal resampling of 3D+t TEE images acquired on multiple heartbeats was proposed to improve temporal consistency [167].

An increased holistic view of the heart, demanded by clinicians is in perfect accordance with the tremendous scientific effort worldwide, such as the Virtual Physiological Human project [30], geared towards multi-scale physiological modeling and simulation, which will promote personalized, preventive and predictive healthcare. However, the majority of cardiac models to date focus on representation of the left or right ventricle [54], the left and right atrium [205], while few model the left side valves but none explicitly handles the entire valve apparatus including the right side valves. A critical component for patient-specific computational models of the entire heart and realistic cardiovascular simulations, which was not reported yet in the literature, is a personalized and complete representation of the valvular apparatus. This would allow for personalized bio-mechanical simulations using patient-specific geometry.

Interactive segmentation methods are a popular technique in medical image analysis where the user provide some information, such as a seed region or rough outline of the region to segment. An algorithm can then iteratively refine such a segmentation, with or without guidance from the clinician [64, 36].

2.7.2. Intra-operative guidance/Registration

Delivering high-quality pre-operative data, or anatomical models derived from the images using segmentation algorithms such as 2.7.1, into the operating room to guide cardiac therapy is highly desirable for minimally invasive procedures, where the target anatomy is not directly visible in the intra-operative images. Thus overlays of 3D anatomical structures based on pre-operative data can provide valuable information for interventional navigation and guidance. High-quality pre-operative 3D information is routinely acquired for diagnostic and planning purposes by means of Computed Tomography, Magnetic Resonance Imaging (MRI) or Echocardiography. Aligning these with intra-operative data is a key goal in interventional guidance. Registration algorithms are an ideal tool to achieve this goal.

Registration is a process that searches for the correct alignment of multiple images and bringing them into the same reference coordinate system. In clinical practice it is used to combine complementary information from multiple images, evaluate temporal changes in longitudinal studies or characterize a population group. The registration approaches can be classified into mono-modal, multi-modal, point or feature based, intensity based, rigid, deformable, pair-wise and simultaneous registration algorithms. For detailed treatment of the subject we refer to chapter 4, reviews and surveys in literature [125, 92, 151, 126, 137].

Within the Hybrid operating rooms modern C-arm systems can acquire fluoroscopy (2D) and 3D C-arm CT (3D) images. Many groups use 3D-2D registration algorithms in order to align the pre-operative image with the intra-operative setting. However, direct 3D pre-operative to 2D fluoroscopy image registration is difficult to solve, especially within the intra-operative setup that does not allow for user interaction or time consuming processing. Linte [115] proposed to use pre-operative CT images and extract models which will be registered with intra-operative images. Major limitations are the required tracking equipment and the semi-automatic delineation of the mitral annulus. Other methods rely on fiduciary markers to achieve 3D/2D registration. Lang [44, 115, 108, 117, 150] proposed a real-time approach to fuse TEE and fluoroscopy by placing markers on the TEE probe.

C-arm CT [110] is emerging as a novel imaging modality that can acquire 3D CT-like volumes directly in the operating room in the same coordinate system as the 2D fluoroscopy images, which overcomes the need for 2D/3D registration [45, 146]. For most procedures, the patients are older and the added radiation compared to fluoroscopy is not a major concern. Instead, a safe and successful execution of the procedure is the dominating factor [122].

Some methods work directly on the C-arm CT images [206] to extract patient-specific models and overlays for procedure guidance, eliminating the need for pre- and intra-operative image fusion completely. However, acquiring high-quality, contrasted, and motion compensated (using rapid-pacing) C-arm CT images is not feasible for all patients. Instead, a much simpler protocol, which acquires non-contrasted, non-ECG-gated C-arm CT volumes can be performed to serve as a bridge between 3D pre-operative images and 2D live fluoroscopy.

Multi-modal 3D/3D registration algorithms can be utilized to align the pre-operative image with the C-arm CT volume. In Wells et al. [200, 191], mutual information is used to cope with intensity inconsistencies between CT and MR. The authors of [212] proposed a novel similarity metric, which incorporates prior knowledge of previously registered images. In [119], an atlas-based approach was presented to track the myocardium and left and right ventricles from MR data. The registration is used to align the cardiac atlas to the patient data. However, these methods are computationally expensive, and without the appropriate guidance of a shape prior likely to converge into local minima.

2.7.3. Computational Modeling

Simultaneously researchers are developing detailed computational models of valve biomechanics driven by the growing prevalence of valvular heart disease. This tools would enamble the next generation of clincal application for advanced intervention planning.

Quantifying valve function, by computing measurments from geometrical models extracted from non-invasive imaging (see section 2.7.1), might not be sufficient to plan the optimal treatment for a specific patient. Mechanical insights are necessary to predict how pathological valve dynamics will be modified after intervention. Furthermore, a comprehensive understanding of the valve physiology is crucial in order to design long-term treatments that do not alter normal valve and heart function. To address these questions, several computational models of valve physiology have been proposed. Since the pioneering work of Kunzelman et al. [107], several models have been proposed and new insights on the valve function have been obtained. Two categories of computational valve models can be distinguished: structural models and fluid-structure interaction (FSI) models. Structural models aim to simulate the biomechanics of MV apparatus without directly considering the blood that flows across it. The standard approach is to use finite-element models (FEM) to solve the dynamics equation of valve anatomy, surface pressure and boundary conditions [107, 106].

Especially in the context of minimally invasive procedures such as Transcatheter Aortic Valve Implantation (TAVI) [194] or the Mitra Clip procedure. During intervention planning for minimally invasive procedure simulating different procedural options such as different device types, sizes and implant position can help the clinician to select the optimal treatment for a particular patient. For the mitral valve several constitutive laws of mitral leaflets have been proposed to simulate the mitral valve biomechanical properties, from simple isotropic linear elasticity to more complex anisotropic and non-linear hyperelasticity [153, 163, 193, 104, 154, 153, 193, 172, 1]. Fluid-structure interactions have also been investigated to study the impact of the blood flow on valve closure [43]. Most of these models have been developed on synthetic or *ex-vivo* anatomies. Patient-specific anatomies and boundary conditions are starting to be used [173] but they require tedious manual delineations from medical images, which makes the necessary large-scale validations in patients very difficult. Recently, a first patient-specific simulation of MV annuloplasty has been presented [173]. In respect to the TAVI procedure, the authors in [196], computed

TAVI deployment on a patient-specific anatomy calculated from CT images. They obtained promising results with a non-linear, hyper-elastic model of aortic apparatus. In [24], finite element modeling was employed to calculate the stresses on an Edwards Sapien implant to predict stent rupture. Similarly, in [181], the authors investigated the radial force generated by Medtronic - CoreValve and Edwards Sapien device.

2.8. Conclusion

The heart is a complex organ which pumps the blood throughout the blood vessels by repeated, rhythmic contractions. Its essential components are the heart valves which regulate the blood flow during the cardiac cycle. Dysfunctions on each of the heart valve have severe implications on the patients health. In fact valvular heart disease is the largest subgroup among all cardiovasclular diseases. Despite advancements in medication and interventional valve therapy remains the most expensive and riskiest among all cardiovascular procedures. Within the last decade novel minimally invasive treatments have emerged. They have the potential to reduce procedural morbidity, mortality, and costs of surgical valve replacement or repair while accelerating patient recovery. However as there is no direct view or access to the affected anatomy pre-operative planning in combination with procedural guidance is crucial for a successful outcome. During the last decade each medical imaging modality has improved with respect to the acquired image quality, as well as the accessibility and cost effectiveness. Techniques like Transesophageal Echocardiography (TEE), cardiac Computed Tomography (CT) and Cardiovascular Magnetic Resonance (CMR) imaging, enable now dynamic four dimensional scanning of the beating heart. In addition modern C-arm systems are capable to acquire a three dimensional CT like modality within the operating room. Thus non-invasive imaging in combination with fast, precise and reproducible image analysis tools will become an essential part of clinical practice to assure a successful outcome. In this context, this thesis offers a novel modeling and guidance paradigm of minimally invasive procedures, which aims to consolidate medical knowledge about the valvular apparatus and substantially benefit the entire clinical management of valvular heart disease patients.

3. Quantification and pre-operative modeling

In the context of treatment of valvular heart disease, quantification and treatment planning are crucial parts of the clinical workflow. In most cases diagnostic and non-invasive volumetric imaging data is available to support the diagnosis and pre-operative planning. Complex patient-specific models extracted from the images can be used to extract precise and reproducible measurement used during intervention planning and quantification of valve morphology and function. Our goal is to automate this process and extract anatomical information from the volumetric data without user interaction. This technique can advance the current clinical workflow of interventional planning, improve procedural outcome, while reducing interventional risk.

Furthermore a novel volumetric model for the aortic valve will be presented in section 3.6 necessary for advanced planning for the TAVI procedure. We thus concentrated on a patient cohort with severely stenotic aortic valves. Tissues within the aortic valve can be classified in one of the following categories: leaflet tissue, blood pool and calcification. In addition to intensity based features we incorporated novel geometrical features to capture the spatial context of each tissue with respect to the aortic valve anatomy. Based on the volumetric model, a computational modeling framework is developed to simulate the implant deployment procedure for TAVI [71]. From a diagnostic CT a geometrical model can be extracted and used to personalize the bio-mechanical computational model.

Finally we propose an algorithm to extract accurate valve models from low-quality contrasted intra-operative images by using additional pre-operative information [69]. Hereby a learning based approach is utilized where both models of the same patient are estimated at once.

3.1. Physiological Model of the Heart Valves

In this section we introduce the complete heart valves model, which includes the aortic, mitral, tricuspid and pulmonary valves, and captures their morphological, functional and pathological variations. To reduce anatomical complexity and facilitate effective estimation, the heart valve model is represented on three abstraction layers [66, 68, 71, 67, 72] :

- **Global Motion Model**: which represents the global location and motion of each valve.
- **Anatomical Landmark Model**: representing the motion of the corresponding anatomic landmarks.

- **Complete Valve Model**: which parameterizes the full anatomy and dynamics of the valves using dense surface meshes.

3.1.1. Global Motion Model

The global dynamic variation of each valve is parameterized through a similarity transformation in the Euclidean three-dimensional space, which includes nine parameters.

$$\boldsymbol{B}_t = \{(c_x, c_y, c_z), (\alpha_x, \alpha_y, \alpha_z), (s_x, s_y, s_z)\} \quad t \in 1 \dots T \tag{3.1}$$

(c_x, c_y, c_z) is the translation, $(\alpha_x, \alpha_y, \alpha_z)$ the quaternion representation of the rotation, (s_1, s_2, s_3) the similarity transform scaling factors and the time variable t is capturing the temporal variation during the cardiac cycle.

3.1.2. Anatomical Landmark Model

A set of 33 anatomical landmarks, described in the next paragraph, are used to parametrize the complex and synchronized motion pattern of all valves, which explains the nonlinearities of the hemodynamic movements. Thereby, each landmark is described by a time-step trajectory T in a three dimensional space, normalized by the temporal dependent similarity transform $\boldsymbol{B}$:

$$\boldsymbol{L}_n(\boldsymbol{B}) = \{l_1,\ l_2,\ \dots l_T\} \quad n \in \{1 \dots 33\} \quad l_i \in \mathbb{R}^3 \tag{3.2}$$

3.1.3. Complete Valve Model

The final valves model is completed with a set of 13 dense surface meshes. Each mesh is sampled along anatomical grids of vertices defined by the landmarks:

$$\boldsymbol{V}_q(\boldsymbol{L}, \boldsymbol{B}) = \{\vec{v_1}, \vec{v_2}, \cdots, \vec{v_K}\} \quad q \in \{1 \dots 13\} \quad \vec{v_i} \in \mathbb{R}^3 \tag{3.3}$$

where $\vec{v_i}$ are the vertices, and K is the total number of vertices of mesh q. Each anatomical landmark has a fixed correspondence on the parametrized surface mesh.

Aortic valve

Four surface structures represent the aortic valve: aortic root, left coronary leaflet, right coronary leaflet and non coronary leaflet. The aortic root connects the ascending aorta to the left ventricle outflow tract and is represented through a tubular grid (see Figure 3.1). This is aligned with the aortic circumferential u and ascending directions v and includes 36×20 vertices and 1368 faces. The root is constrained by six anatomical landmarks, i.e. three commissures and three hinges, with a fixed correspondence on the grid. The three aortic leaflets, the L-, R- and N-leaflet, are modeled as paraboloids on a grid of 11×7 vertices and 120 faces (see Figure 3.5(c)). They are stitched to the root on a crown like

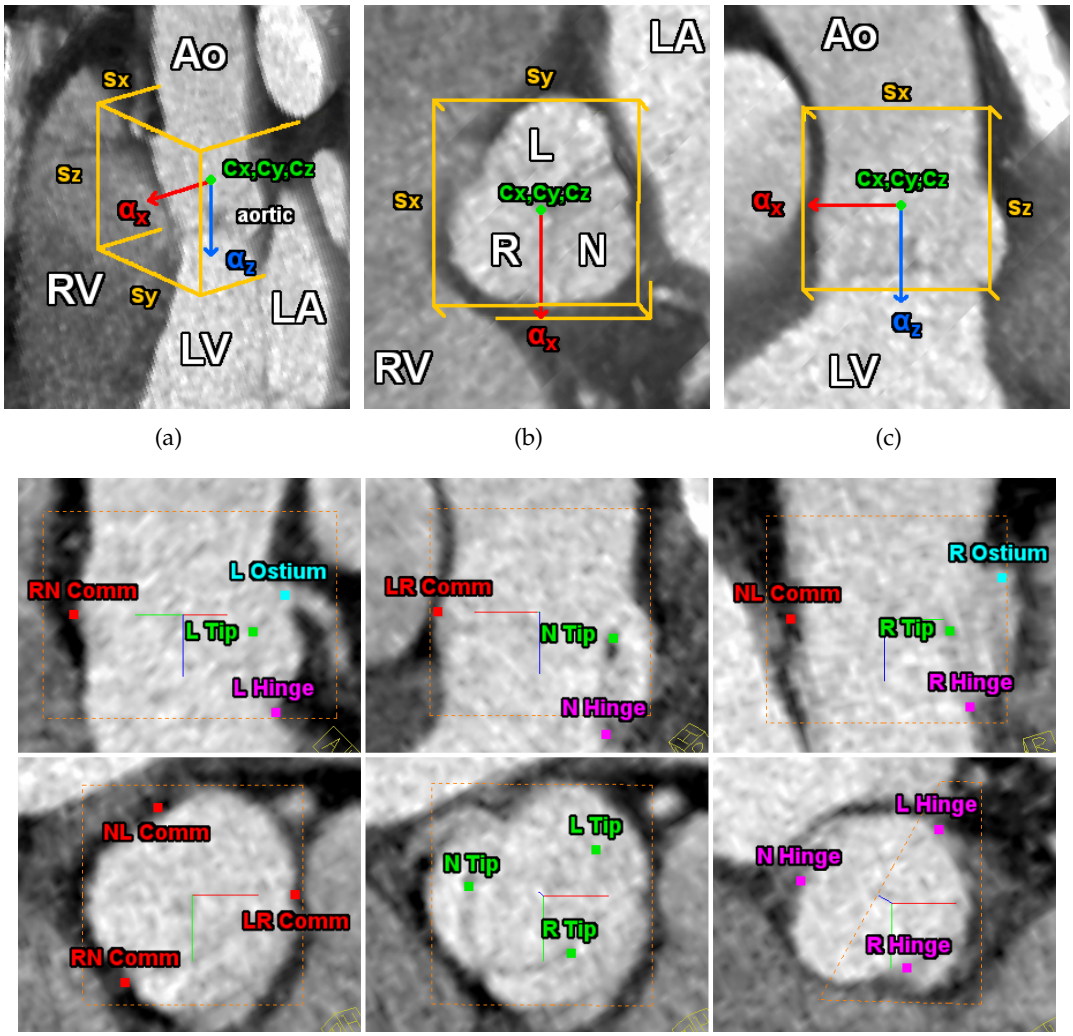
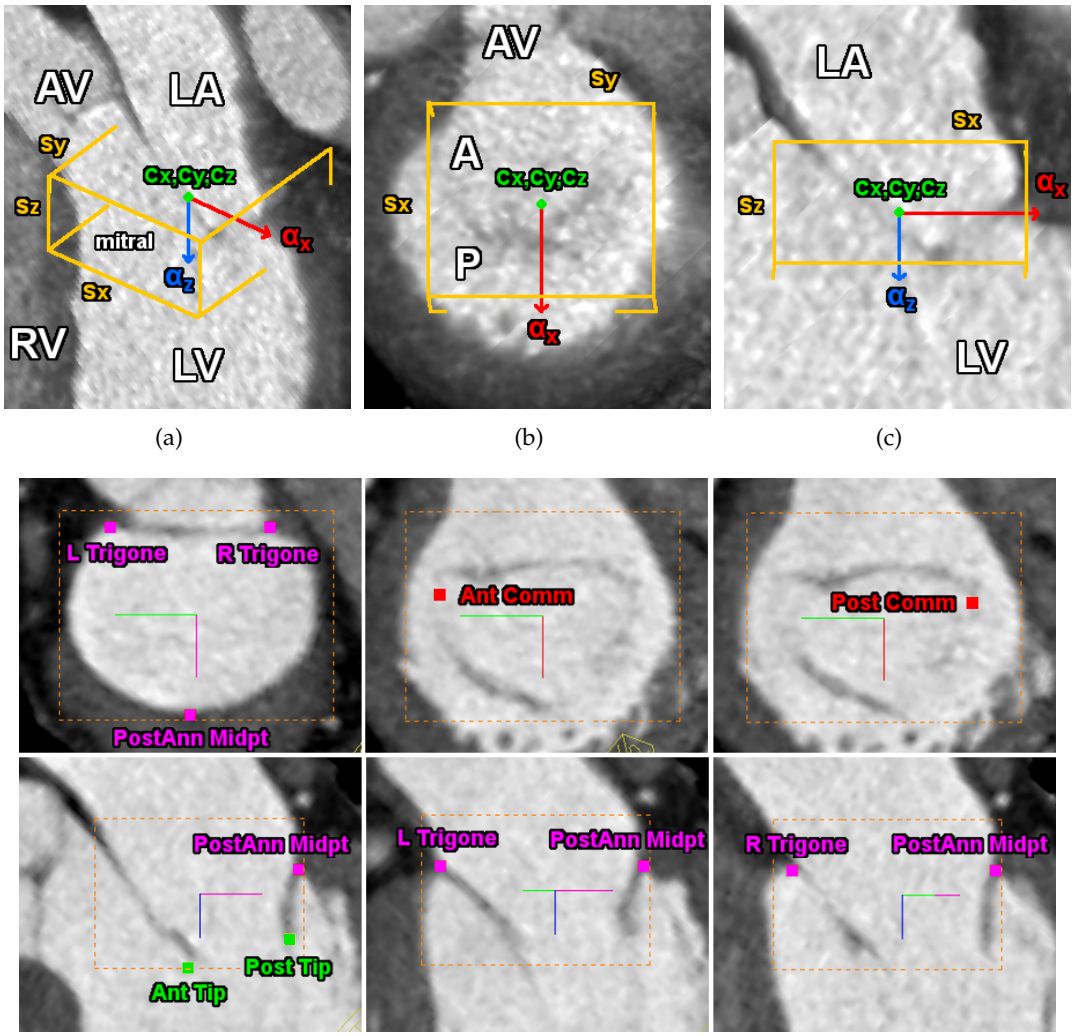


Figure 3.1.: Global motion and anatomical landmark model of the aortic valve. The similarity transform is represented as a bounding box around the aortic valve estimated from 4D cardiac CT. (a) Perspective view; (b) Long Axis; (c) Short Axis; (d) Landmarks relative to the anatomical location illustrated in long and short axis from an example CT study.

attachment ring, which defines the parametric u direction at the borders. The vertex correspondence between the root and leaflets along the merging curve is symmetric and kept fixed. The leaflets are constrained by the corresponding hinges, commissures and tip landmarks, where the v direction is the ascending vector from the hinge to the tip.



(a) (b) (c)

(d)

Figure 3.2.: Global motion and anatomical landmark model of the mitral valve. The similarity transform is represented as a bounding box around the mitral valve estimated from 4D cardiac CT. (a) Perspective view; (b) Long Axis; (c) Short Axis; (d) Landmarks relative to the anatomical location illustrated in long and short axis from an example CT study.

Mitral valve

The mitral valve is composed of 7 landmarks including 3 trigons, 2 commissures and 2 leaflet tips (see Figure 3.2). The leaflets separate the left atrium and left ventricle hemodynamically and are connected to the endocardial wall by the saddle shaped mitral annulus. Both are modeled as paraboloids and their upper margins implicitly define the annulus. Their grids are aligned with the circumferential annulus direction u and the orthogonal

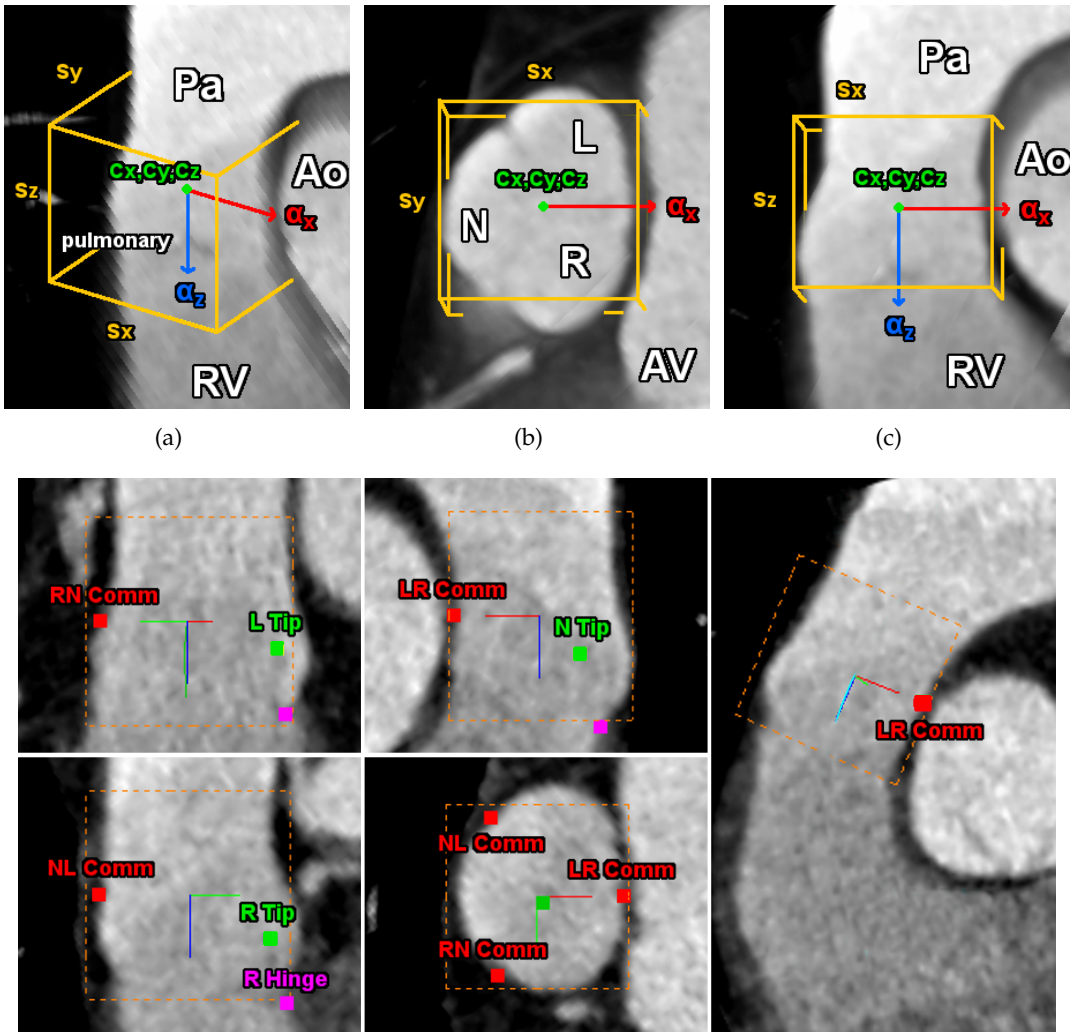
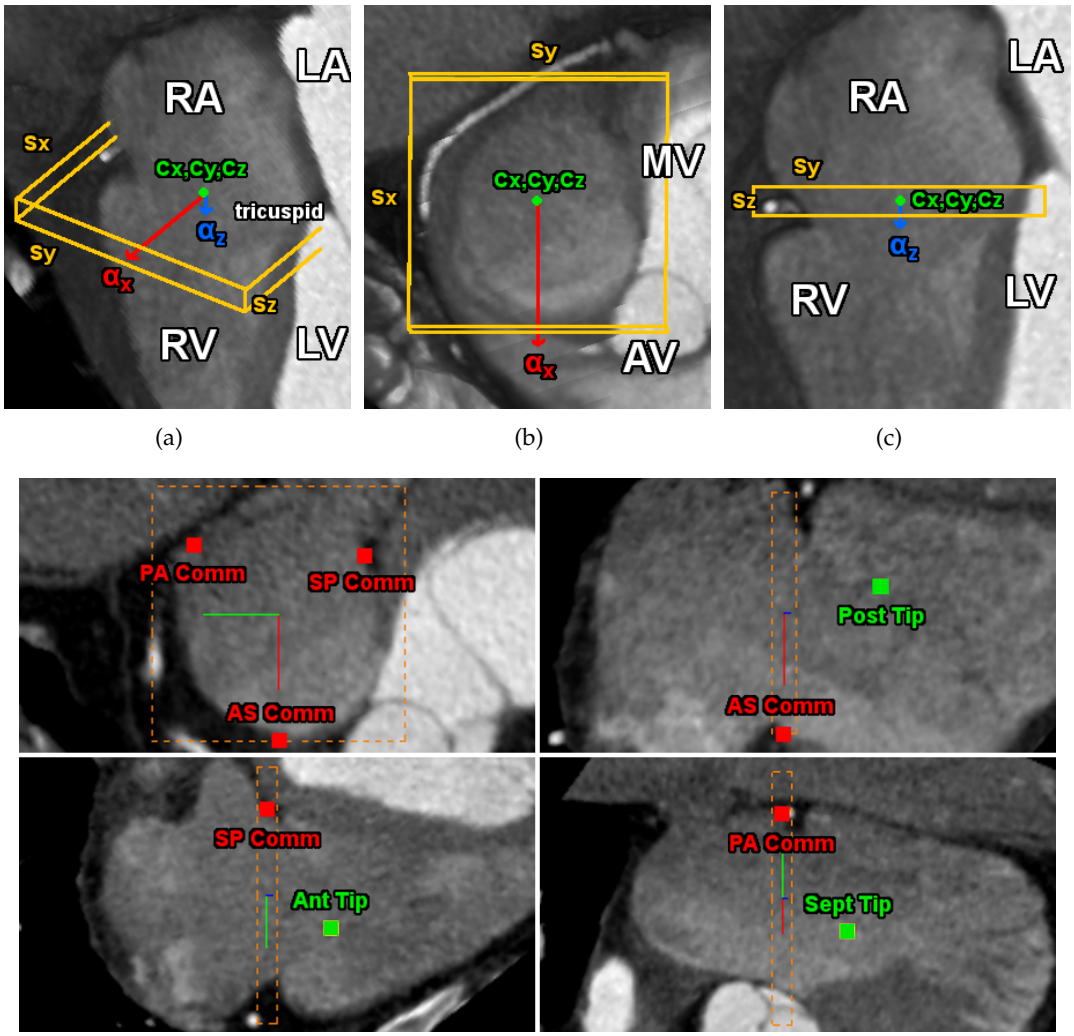


Figure 3.3.: Global motion and anatomical landmark model of the pulmonary valve. The similarity transform is represented as a bounding box around the pulmonary valve estimated from 4D cardiac CT. (a) Perspective view; (b) Long Axis; (c) Short Axis; (d) Landmarks relative to the anatomical location illustrated in long and short axis from an example CT study.

direction v pointing from the annulus towards leaflet tips and commissures (see Figures 3.5(b) and 3.5(d)). The anterior leaflet is constructed from 18×9 vertices and 272 faces while the posterior leaflet is represented with 24×9 vertices and 368 faces. Both leaflets are fixed by the mitral commissures and their corresponding leaflet tips. The left / right trigons and the postero-annular midpoint further confine the anterior and posterior leaflets, respectively.



(a) (b) (c)

(d)

Figure 3.4.: Global motion and anatomical landmark model of the tricuspid valve. The similarity transform is represented as a bounding box around the tricuspid valve estimated from 4D cardiac CT. (a) Perspective view; (b) Long Axis; (c) Short Axis; (d) Landmarks relative to the anatomical location illustrated in long and short axis from an example CT study.

Pulmonary valve

The representation of the pulmonary valve is compounded out of four structures: pulmonary trunk, left facing leaflet, none facing leaflet and right facing leaflet (see Figure 3.3). The pulmonary trunk emerges out of the right ventricular outflow tract, supports the pulmonary valves and its three leaflets and ends at the level of the pulmonary artery bifurcation. The grid, which spans the pulmonary trunk surface, is aligned with the cir-

cumferential u and longitudinal direction v of the valve. It includes 50×40 vertices and 3822 faces confined through the pulmonary commissures, hinges and the RV trigon. The attached L-, R- and N- leaflets, are modeled as paraboloids along the annulus circumferential direction u and vector v pointing from the corresponding hinge to the leaflet tip (see figure 3.6(c)). Each includes 11×7 vertices and 120 faces bounded by the associated two commissures, hinge and tip.

Tricuspid valve

The function of the tricuspid valve is to regulate the blood flow from the right atrium to the right ventricle, staying closed during systole and opened during diastole. The model is constrained by four surface geometries (annulus, septal-, anterior- and posterior leaflet) (see Figure 3.6(d)) and six anatomical landmarks (three commissures and three leaflet tips as illustrated in Figure 3.4) which are corresponding to vertices on the meshes. The tricuspid annulus is represented as a surface mesh constrained by the three commissures.

The tricuspid leaflets are modeled as hyperbolic paraboloids and implicitly describe the tricuspid annulus. Their grids are spanning along the annulus circumferential direction u and the perpendicular vector v pointing for the annulus towards the corresponding leaflet tip, and consist out of 22×14 vertices and 546 faces. Each leaflet is constrained by the corresponding two commissures and one leaflet tip (see Figure 3.6(d)) .

3.2. Discriminative machine learning techniques

Estimating the parameter of the complete heart valve model is a challenging task. Due to the large number of parameters within the complete valvular model $\boldsymbol{V}_q(\boldsymbol{L}, \boldsymbol{B})$, the complex morphology and function of the valves and the limiting imaging quality and resolution, robust tools must be employed. Given a large quantity of training data is available, supervised discriminative machine learning techniques have been shown to be very effective in this setting. In many areas of computer vision such as object detection, categorization and segmentation methods based on supervised discriminative machine learning techniques are consistently outperforming other methods.

Based on a volumetric image I, the discriminative model can be formulated in a probabilistic framework to approximate the posterior probability $P(\boldsymbol{V}_q(\boldsymbol{L}, \boldsymbol{B}) \,|\, I)$. In contrast to generative models which require to model the joint probability distribution across the parameter space, discriminative models approximate the conditional probability distribution. Even though generative models have several advantages such as a higher flexibility to model and express complex relationships between observer variables and target variables their performance on classification tasks is usually worse than discriminative models [95, 90, 16]. In addition in many cases it is not possible to model the joint probability across the high-dimensional parameter space and thus discriminative approaches are the only viable alternative.

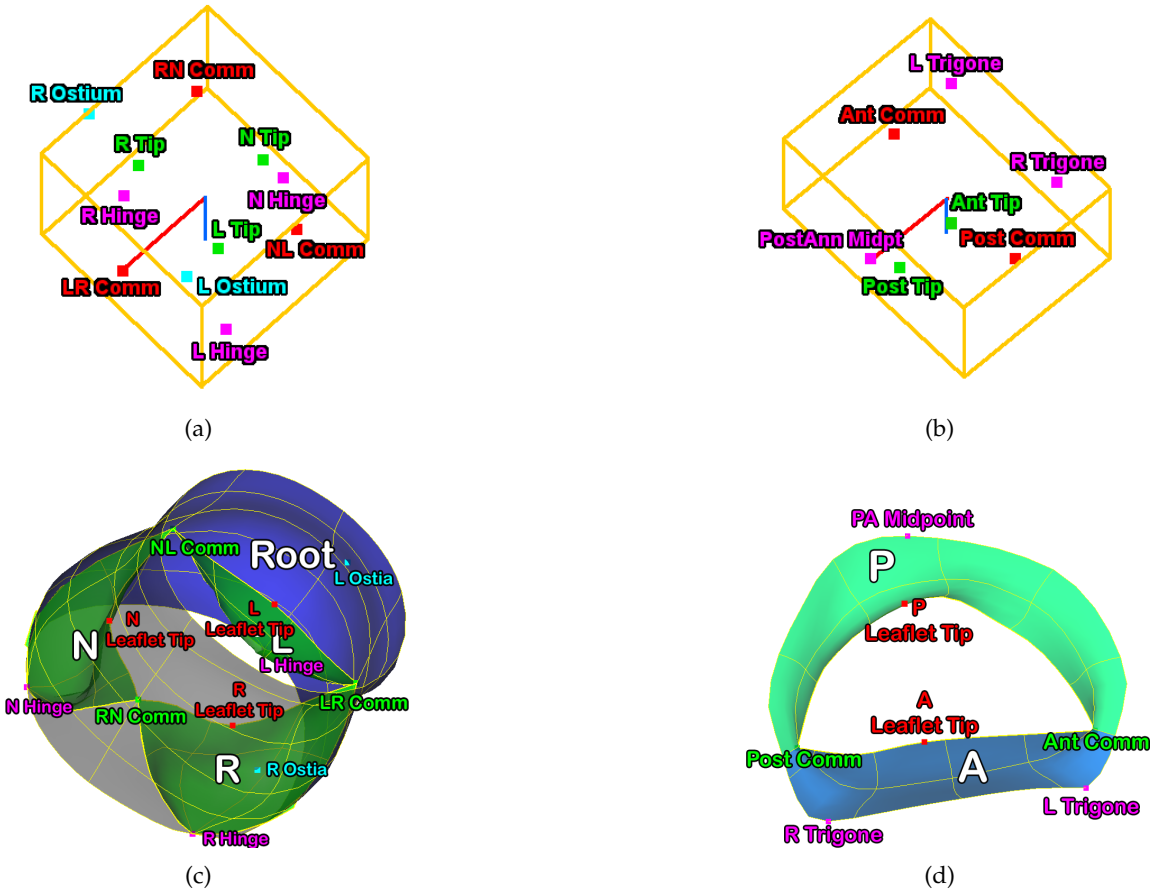


Figure 3.5.: Anatomical Landmark Model and Complete Valve Model of the aortic valve, mitral valve, pulmonary valve and tricuspid valve.

Within our supervised approach an annotation database is used to train a discriminative classifier (boosting, support vector machines, randomized trees) on a set of image features. The goal of the classifier is to discriminate between positive and negative examples. Throughout our work we use Boosting as the classification technique. In section 3.2.1 we explain the main properties of Boosting as all of our methods described in this thesis are based on this framework. Further the image features which are used as input to the boosting classifier are described in section 3.2.2.

3.2.1. Boosting

Boosting is a supervised machine learning algorithm. The main idea is to compose a strong classifier from a group of weak classifier. The only restriction of the weak classifier performance is that the classification result has to be slightly better than random guessing. However as the weak learners are seen as complementary, when combined into the strong

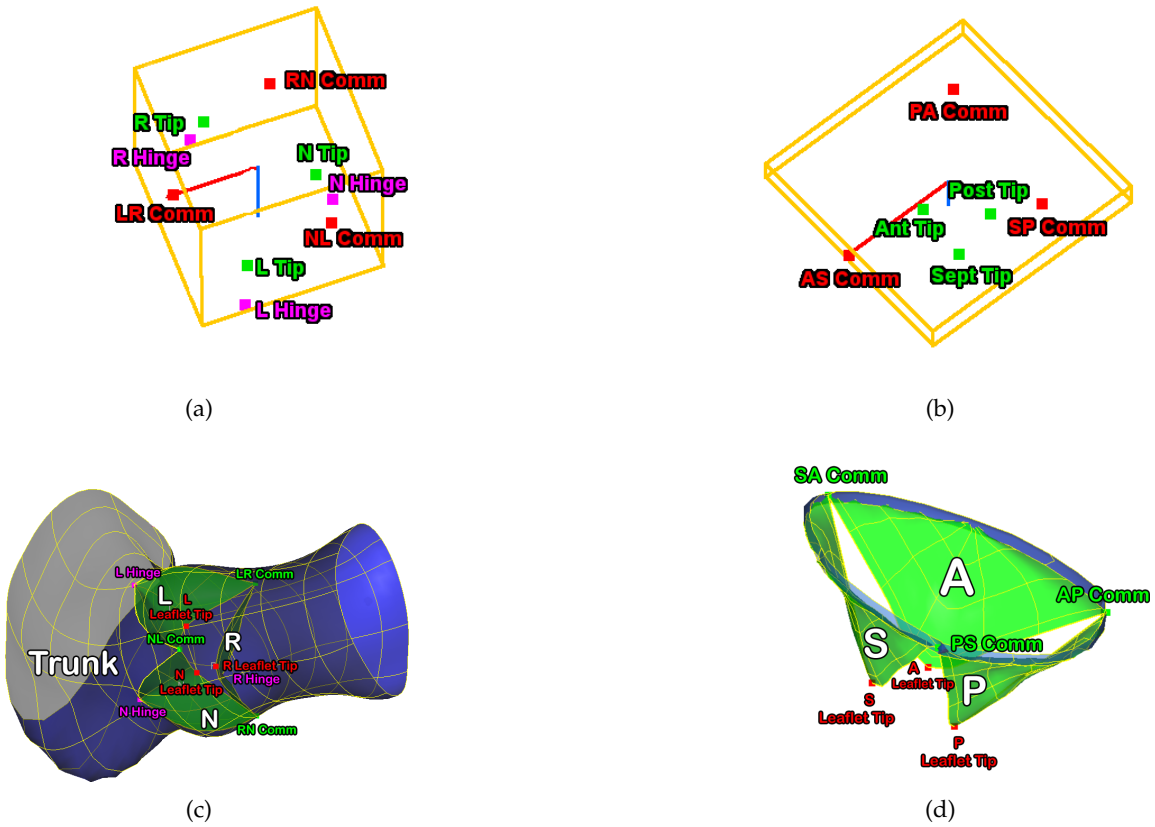


Figure 3.6.: Anatomical Landmark Model and Complete Valve Model of the aortic valve, mitral valve, pulmonary valve and tricuspid valve.

classifier the classification is well-correlated with the true classification task.

Based on a training set $\{(x_1, y_1), ..., (x_N, y_N)\}$, where x_i belongs to a certain domain X, and y_i is a binary label set $Y = \{-1, 1\}$ the boosting algorithm selects a group of weak learners $\{h_1, ..., h_l\}$, with $h_i : X \rightarrow Y$, in order to generate a strong classifier $D(x)$.

While the boosting framework is generic, most algorithms realizing the principle of boosting share the following properties. An iterative process is used to select the weak learners. When they are selected, they are re-weighted according to the weak learners accuracy in respect to the classification task. After a weak learner is selected, the training data is re-weighted: examples that are misclassified gain weight while examples that are classified correctly lose weight.

The first version of the algorithm was proposed by Robert Schapire and Yoav Freund [51]. However the weights of the weak learners were not adaptive. Current boosting techniques differentiate in the way they re-weight training data and the weak learners. AdaBoost is the most popular as it was the first algorithm that could adapt the weak learners [162]. However, there are many more recent algorithms such as LPBoost [197], To-

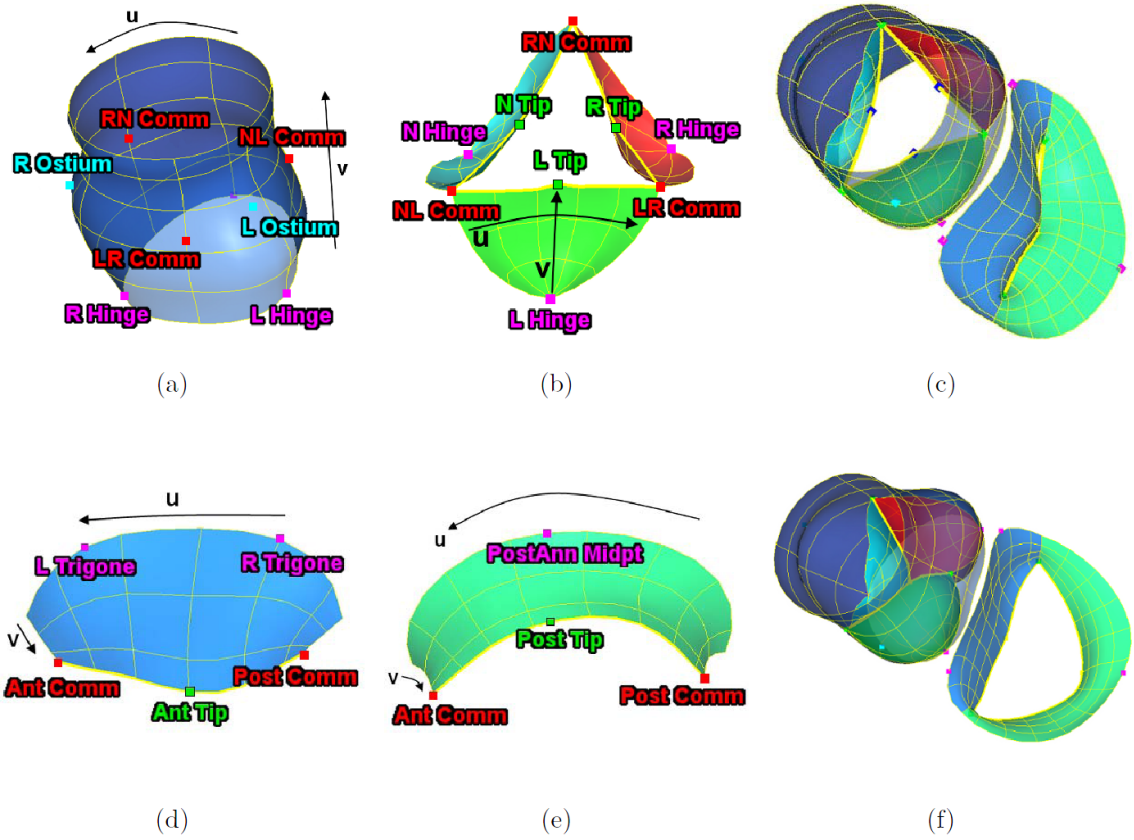


Figure 3.7.: Isolated surface components of the aortic and mitral models with parametric directions and spatial relations to anatomical landmarks: (a) aortic root, (b) aortic leaflets, (c) aortic-mitral in end-systole, (d) anterior mitral leaflet, (e) posterior mitral leaflet and (f) aortic-mitral in end-diastole.

talBoost [198], BrownBoost, MadaBoost [41], LogitBoost [53], and others. Many boosting algorithms fit into the AnyBoost framework, which shows that boosting performs gradient descent in function space using a convex cost function.

In the next two chapter the Adaboost and the Probabilistic Boosting Tree (Probabilistic Boosting Tree (PBT)) will be presented. Throughout the thesis the PBT will be used as the main classification technique.

AdaBoost

AdaBoost, short for Adaptive Boosting, was first presented by Yoav Freud and Robert Shapire [51]. During training AdaBoost adapts the reweighing schema to both the weak learner and the input training data. Within the iterative approach weak learner are selected based on their ability to perform the classification task better than random guessing. Even if the weak learner is worse than random guessing by negative re-weighting it can still be

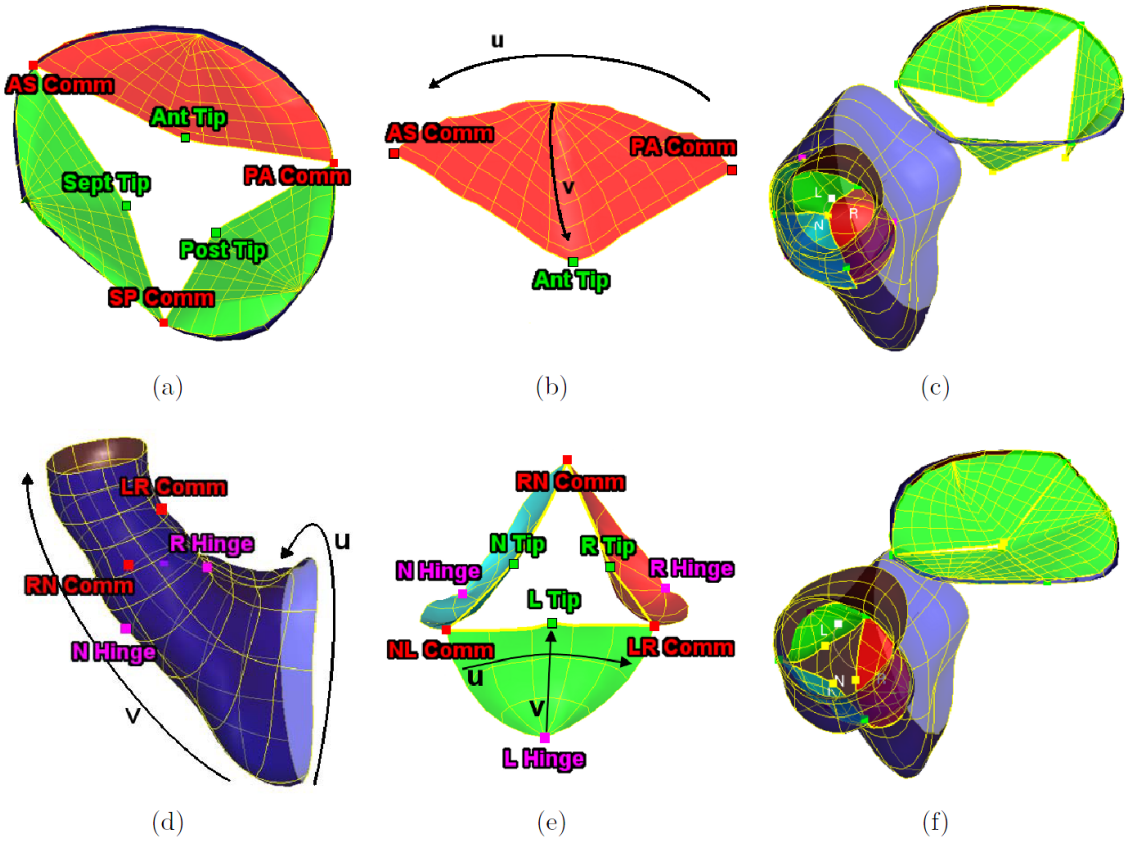


Figure 3.8.: Isolated surface components of the tricuspid and pulmonary models with parametric directions and spatial relations to anatomical landmarks: (a) tricuspid leaflet, (b) tricuspid annulus and leaflets, (c) tricuspid-pulmonary in end-diastole, (d) pulmonary trunk, (e) pulmonary leaflets and (f) tricuspid-pulmonary in end-systole.

adopted. AdaBoost generates a new weak classifier in each of a series of T rounds . In each round, a distribution of weights is updated that indicates the importance of training examples in the data set for the classification. The weights of each incorrectly classified example are increased, and the weights of each correctly classified example are decreased. Thus in the next iteration the new learner focuses on the hard examples. The algorithm is shown in Figure 3.9.

It is shown that AdaBoost is less susceptible to the over-fitting problem than most learning algorithms and thus shows a good generalization performance. In addition the number of parameters is small compared to other algorithms. In addition the scaling of the input space is not important. Thus in contrast to other learning algorithms the individual dimensions of the input space do not need to be scaled.

Input

- m labeled samples $\{(x_1, y_1), ..., (x_N, y_N)\}$ where $x_i \in X$, $y_i \in Y = \{-1,\ +1\}$
- distribution D over the examples
- weak learning algorithm producing weak hypothesis h
- number of iterations $1 \ldots T$

Initialization

- normalize the distribution D over m samples $D_1(i) = 1/m$ for $i = 1, ..., m$

Main Loop, for t=1,...,T

- Train weak learner using sample distribution D_t
- Get weak hypothesis $h_t :\ X \to \{0,\ +1\}$ with error $\epsilon_t = Pi_{D_{t(i)}}\left[h_t(x_i) \neq y_i\right]$
- Choose $\alpha_t = \frac{1}{2}\ln\left(\frac{1-\epsilon_t}{\epsilon_t}\right)$
- Update:

$$\begin{aligned} D_{t+1}(i) &= \frac{D_t(i)}{Z_t} \times \begin{cases} e^{-\alpha_t} & : \text{ if } h_t(x_i) = y_1 \\ e^{-\alpha_t} & : \text{ if } h_t(x_i) = y_1 \end{cases} \\ &= \frac{D_t(i) exp(-\alpha_t y_i h_t(x_i))}{Z_t} \end{aligned}$$

  where Z_t defines the partitioning function (or normalization factor) chosen that D_{t+1} will be a probability distribution.

Output:

- $$H(x) = \text{sign}\left(\sum_{t=1}^{T} \alpha_t h_t(x)\right) \tag{3.4}$$

Figure 3.9.: The adaptive boosting (AdaBoost) algorithm introduced by Freund and Schapire in [51].

Probabilistic Boosting Tree

There are several problems with the AdaBoost algorithm. First, though it asymptotically converges to the target distribution, it may need to pick hundreds of weak classifiers. This happens when the error rate ϵ_t approaches $1/2$. Second, the order in which features are picked in the training stage is not preserved. The order of a set of features may corre-

spond to high-level semantics and, thus, it is very important for the understanding of objects/patterns. Having a tree structured classifier can preserve the feature order.

In contrast the Probabilistic Boosting Tree (PBT) constructs a binary tree structured classifier where each node represents a strong boosting classifier, composed of a group of weak classifiers. Thus instead of putting all the weak classifiers together into a single strong classifier, a divide-and-conquer approach is used to approximate the target posterior distribution by data augmentation (tree expansion). See Figures 3.10, 3.11, 3.12 and 3.13 for more details.

During training the PBT algorithm recursively constructs a binary tree. At each node, a strong classifier is learned using the standard boosting algorithm. The training samples are divided into two new sets using the learned classifier, the left and the right, which are used to train a left sub-tree and right sub-tree respectively. The variable ϵ_t is used to control the over-fitting problem. Samples falling within the range of $(\frac{1}{2} - \epsilon_t; \frac{1}{2} + \epsilon_t)$ are confusing ones and will be used in both the left and the right sub-trees for training. By reducing the number of input samples the complexity of the problem is reduced as well which leads to a better decision boundary. Thus samples are naturally divided into subgroups. Figure 3.11 shows an example of how a tree is learned and the training samples are divided. Samples which are hard to classify are passed down leading to the expansion of the tree. Clustering of positives and negatives is naturally performed. Since each tree node is a strong classifier, it can deal with samples with complex distribution. There is no need to pre-specify the number of clusters. The hierarchical structure of the tree determines the clusters according to different levels of discrimination.

The testing stage is analogous with the training stage. Figure 3.13 gives the details of how to compute the approximated posterior $p(y|x)$. At the root node of the tree, it gathers the information from its children nodes and reports an overall approximated posterior distribution. This algorithm can also be turned into a classifier which makes hard decision. After computing $p(+1|x)$ and $p(-1|x)$, one can decide to go into the right or left sub-trees by comparing $p(+1|x)$ and $p(-1|x)$. The empirical distribution $q(y|x)$ contained at the leaf node of the tree is then passed back to the top node of the tree. Once a PBT is trained, the $p(+1|x)$ can be used as a threshold to balance between precision and recall. In contrast, a traditional cascade approach needs to train different classifiers based on different precision requirements.

3.2.2. Image Based Features

In the context of classifying objects within medical images the PBT classifier is used in combination with image features. The weak learner within each tree node represents features extracted from the volumetric medical images. From a large pool of image features which are associated with each sample the boosting algorithm selects the most discriminative one to separate the samples in a positive and negative group.

In recent decades several image features have been developed. Most of them were introduced in the area of computer vision used for a specific group of applications which

Input

- A training set $S = \{(x_1, y_1, w_1), ..., (x_N, y_N, w_N)\}; x_i \in \chi, y_i \in \{-1, +1\}, \sum_1 w_i = 1$
- Tree maximum depth L and confusion tolerance ϵ, e.g. $\epsilon = 0.1$
- compute empirical distribution $\widehat{q}(y) = \sum_i w_i \delta(y_i = y)$

Main Loop

- Exit if current tree depth is L
- From the training set S learn a strong classifier using a boosting algorithm with T weak classifiers and early exit $\epsilon_t > \theta$, e.g. $\theta = 0.45$
- Initialize empty sets S_{left} and S_{right}
- For each (x_i, y_i) compute the probability $q(+1|x_i)$ and $q(-1|x_i)$ from the learned strong classifier
- if $q(+1|x_i) - \frac{1}{2} > \epsilon$ then $(x_i, y_i, 1) \rightarrow S_{right}$
- else if $q(-1|x_i) - \frac{1}{2} > \epsilon$ then $(x_i, y_i, 1) \rightarrow S_{left}$
- else $(x_i, y_i, q(+1|x_i)) \rightarrow S_{right}$ and $(x_i, y_i, q(-1|x_i)) \rightarrow S_{left}$
- Normalize all the sample weights in S_{right} and repeat procedure recursively
- Normalize all the sample weights in S_{left} and repeat procedure recursively

Figure 3.10.: The probabilistic boosting-tree training as introduced by Zhuowen Tu in [179]

require specific properties: computation speed, invariance to rotation and scale changes, robustness to noise, discriminative power etc. Features such as Haar-like features were introduced by Viola and Jones [190]. They constructed a boosted cascade of simple classifiers based on Haar-like features that measure vertical, horizontal, central, and diagonal variations of pixel intensities. Later features such as scale-invariant feature transform (SIFT) introduced by David Lowe [120] had many advantages such as orientation, scale invariance and robustness to illumination changes. However the computation time was significant compared to simple intensity based methods. Later multiple methods were introduced to improve the computational performance such as speeded up robust features (SURF) [10] or histogram of gradients (HOG) [39] features. However both still require significant computation time.

In our case speed is a crucial factor. Being able to perform classification within milliseconds is important for many clinical applications, especially in the context of interventional guidance. Thus we use two main types of features with high computational performance: 3D-extensions of the Haar-like features (see 3.2.2) and steerable features (see 3.2.2).

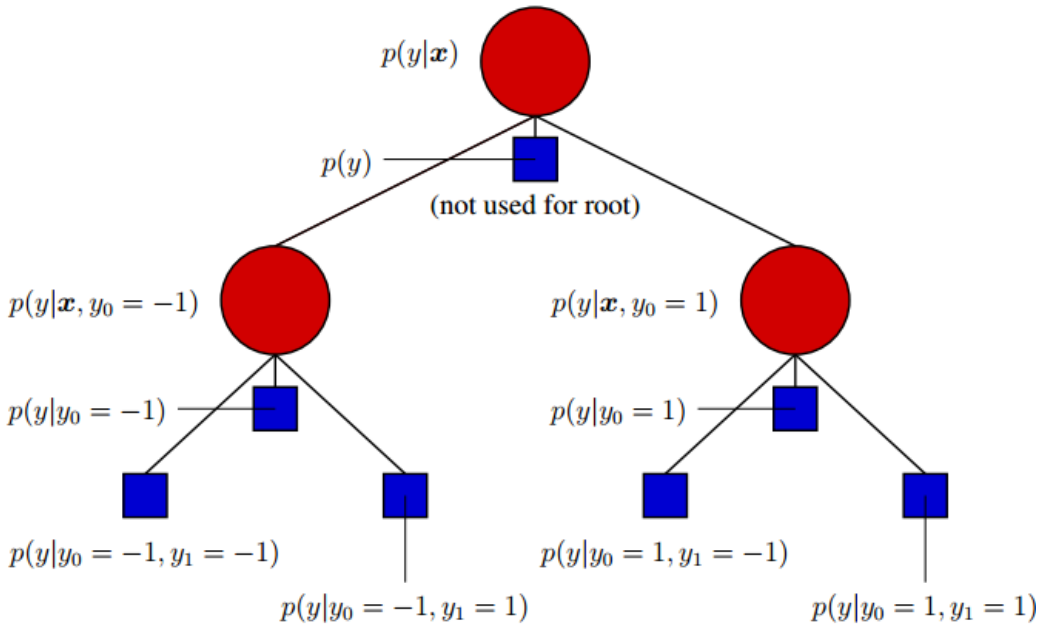


Figure 3.11.: Illustration of the probabilistic model of the tree. The circles are AdaBoost classifiers, the squares contain the empirical class distribution. Each tree node is a strong classifier. Figure from Tu [179].

Haar-Features

Haar-Features owe their name to their intuitive similarity with Haar wavelets. They were first introduced in computer vision for face detection by Viola and Jones [190]. Hereby the main goal was to extract robust features efficiently and use them for object detection. These features are defined as the difference between sums of neighboring image regions, see Figure 3.16.

The sum of image values $i(x', y')$ on a rectangle $(x_0, y_0] \times (x_1, y_1]$ can be computed as:

$$A = \sum_{x_0 < x' \leq x_1} \sum_{y_0 < y' \leq y_1} i(x', y') \tag{3.5}$$

It is computationally expensive, since its complexity depends on the rectangle size. By using the integral image I as an intermediate array A can be computed efficiently (see Figure 3.14 and 3.15). Thus the integral image value at the pixel (x, y) is defined as the sum of the original image values on the rectangle $[0, 0] \times [x, y]$.

$$I(x, y) = \sum_{0 < x' \leq x_1} \sum_{0 < y' \leq y_1} i(x', y') \tag{3.6}$$

The integral image is computed in one pass over the image using the recurrence

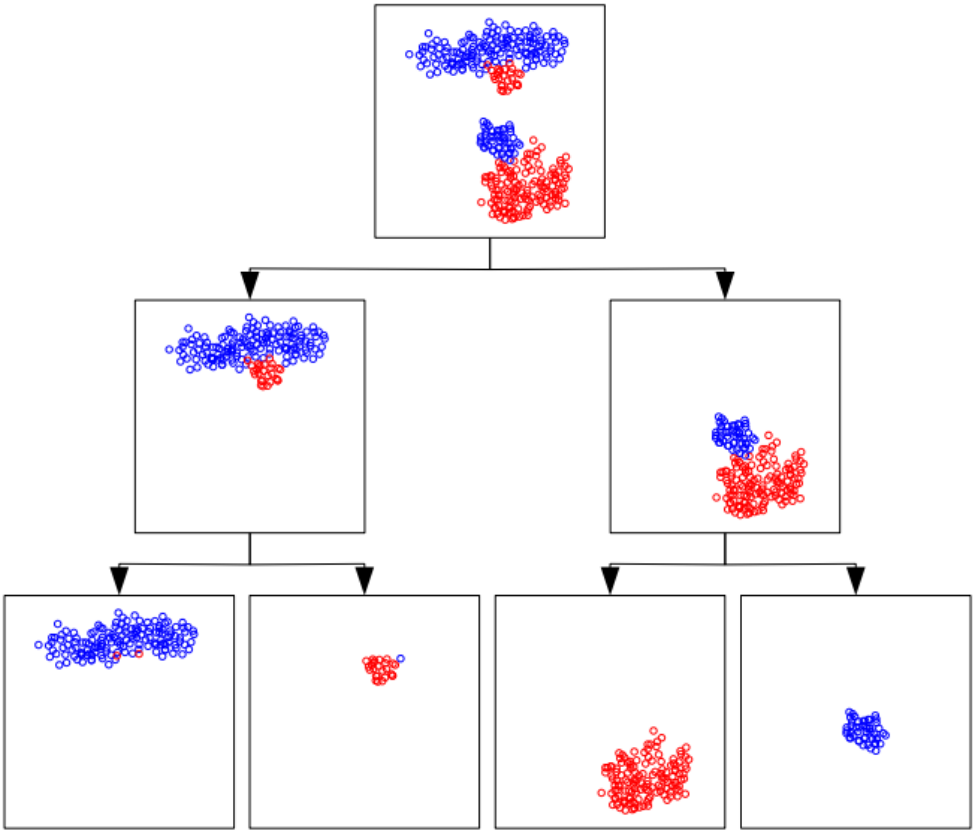


Figure 3.12.: Illustration of how the sample set is separated by the nodes. The first node is trained on all samples. It splits the set into two subsets that may overlap, and the child nodes are trained on these subsets. Red and blue points correspond to the specific class. Figure from Tu [179].

Compute $p_N(y|x)$ the posterior distribution at a tree node N

- Compute $q_N(+1|x)$ and $q_N(-1|x)$ at level N from corresponding strong classifier
- $p_N(y|x) = q(+1|x)p_{right}(y) + q(-1|x)p_{left}(y)$
- If $q(+1|x) - \frac{1}{2} > \epsilon$ then $p_{right}(y) = p_{right(N)}(x, y)$ and $p_{left}(y) = q_{left(N)}(y)$
- Else if $q(y + 1|x) - \frac{1}{2} > \epsilon$ then $p_{right}(y) = q_{right(N)}(y)$ and $p_{left}(y) = p_{left(N)}(x, y)$
- Else $p_{right}(y) = p_{right(N)}(x, y)$ and $p_{left}(y) = p_{left(N)}(x, y)$

Figure 3.13.: The probabilistic boosting-tree testing as introduced by Zhuowen Tu in [179]

$$c(x, y) = c(x, y - 1) + i(x, y) \tag{3.7}$$

$$I(x, y) = I(x - 1; y) + c(x, y) \tag{3.8}$$

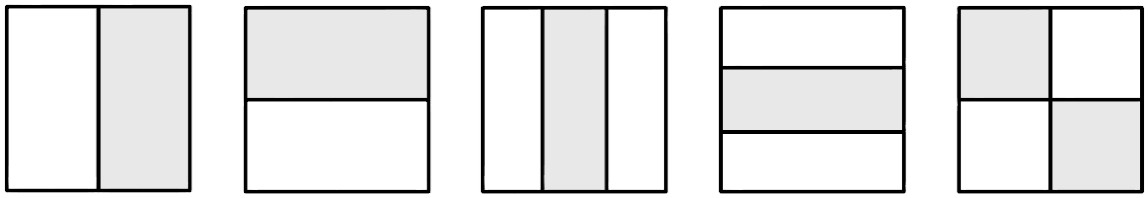


Figure 3.14.: Haar-based rectangular features used for face recognition. The features are the sum on the values on the gray region minus the sum on the white region.

with

$$c(x, -1) = I(-1, y) = 0; \tag{3.9}$$

where $c(x, y)$ is called the cumulative row sum (see Figure 3.15). Thus, one can compute A in constant time using only four references (lookups) to the integral image:

$$A = I(x_1, y_1) - I(x_1, y_0) - I(x_0, y_1) + I(x_0, y_0) \tag{3.10}$$

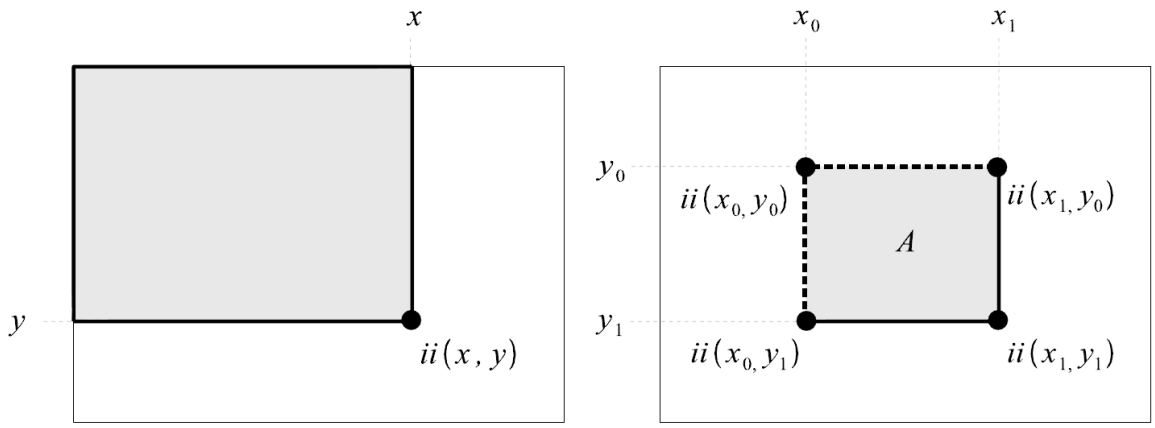


Figure 3.15.: Left: Integral image representation. Right: The four references used to compute the image values on the gray area.

Thus once the integral image is computed a specific haar feature can be computed with a fixed amount of lookups in the integral image. For a sliding window classification this is an efficient way to compute features. This makes them suitable for real-time 2D applications and in the context of medical imaging it offers the ability to detect anatomies in large 3D images within milliseconds.

Steerable-Features

Global features, such as 3-D Haar-like features, are effective to capture the global information (e.g. position, orientation and scale) of an object. To capture the orientation informa-

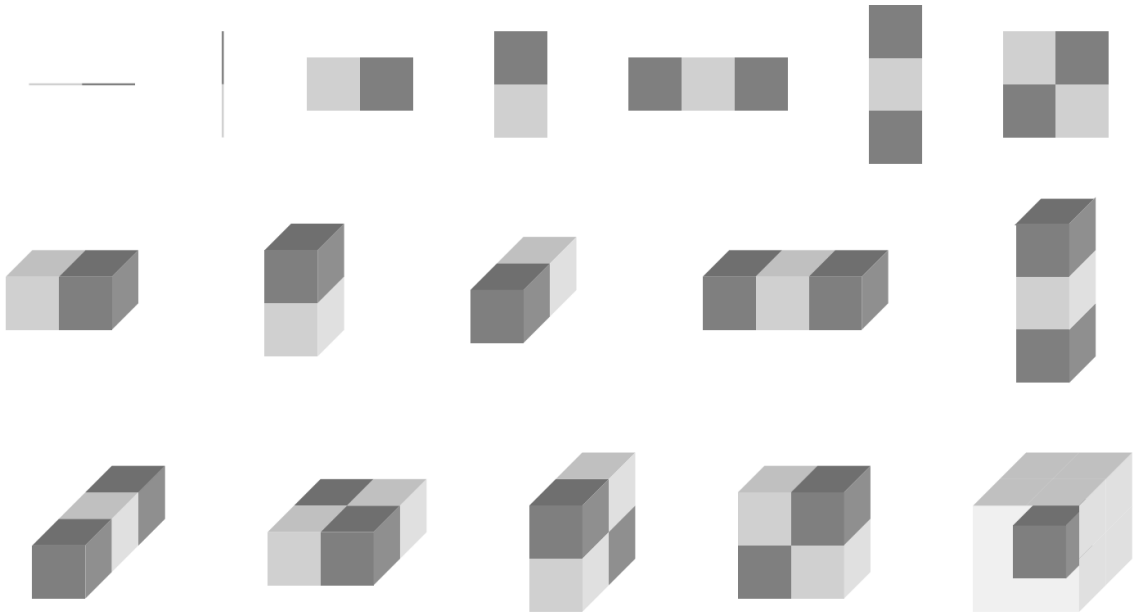


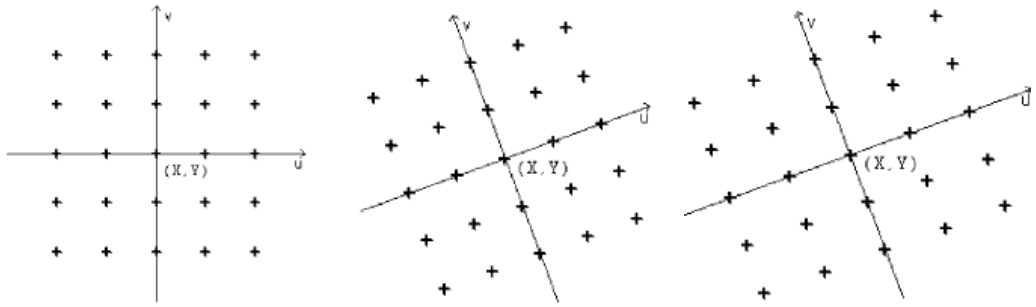
Figure 3.16.: Examples of rectangle 3D Haar features. The sum of the pixels which lie within the white rectangles are subtracted from the sum of pixels in the grey rectangles.

tion of a sample or hypothesis, one would need to either rotate the volume or the feature templates. However, using Haar features this process would be time consuming. In order to test a specific orientation of a sample the complete volume would need to be rotated (re-sampled) and the integral image would need to be recomputed. Steerable features which are defined locally, represent features such as image intensities or local gradients, which are fast to evaluate but lose the global information of the whole object.

Steerable features can capture the orientation and scale of the object while being efficient. In order to compute them a few points are sampled from the volume under a specific sampling pattern (see Figure 3.17). A few local features for each sampling point (e.g., voxel intensity and gradient) are extracted from the original volume. The novelty of our steerable features is that the orientation and scale information is embedded into the distribution of sampling points, while each individual feature is locally defined. Instead of aligning the volume to the hypothesized orientation, the sampling pattern is steered. Figure 3.17 shows how to embed a hypothesis in steerable features using a regular sampling pattern (illustrated for a 2-D case for clearance in visualization). Suppose we want to test if hypothesis $(x, y, z, \vec{\alpha}_x, \vec{\alpha}_y, \vec{\alpha}_z, s_x, s_y, s_z)$ of the similarity transformation of the object is correct. A local coordinate system is defined to be centered at position (x, y, z) (Figure 3.17 a)) and the axes are aligned with the hypothesized orientation $\vec{\alpha}_x, \vec{\alpha}_y, \vec{\alpha}_z$ (Figure 3.17 b)). A few points (represented as + in 3.17) are uniformly sampled along each coordinate axis inside a box. The sampling distance along an axis is proportional to the scale of the shape in

that direction (s_x, s_y, s_z) to incorporate the scale information (Figure 3.17 c)). The steerable features constitute a general framework, in which different sampling patterns [204, 201] can be defined.

At each sampling point, we extract a few local features based on the intensity and gradient from the original volume. A major reason to select these features is that they can be extracted fast. Suppose a sampling point (x, y, z) has intensity I and gradient $g = (g_x, g_y, g_z)$. The tree axes of object-oriented local coordinate system are n_x, n_y and n_z. The angle between the gradient g and the z axis is $\alpha = \arccos(n_z \cdot g)$, where $n_z \cdot g$ means the inner product between two vectors n_z and g. The following 24 features are extracted: I, $\sqrt{I}$,$\sqrt[3]{I}$,I^2,I^3,$logI$,$\|g\|$, $\sqrt{g}$, $\sqrt[3]{g}$, $||g||^2$,$||g||^3$,$log||g||$, α, $\sqrt{\alpha}$, $\sqrt[3]{\alpha}$,α^2,α^3, $log\alpha$, g_x, g_y, g_z, $n_x \cdot g$,$n_y \cdot g$,$n_z \cdot g$. In total, we have 24 local features for each sampling point. The first six features are based on intensity and the remaining 18 features are transformations of gradients.



(a) (b) (c)

Figure 3.17.: Steerable sampling pattern aligned with an example hypothesis $(x, y, \vec{\alpha}_x, s_x, s_y)$ for a two-dimensional problem. Sampling location are defined as '+'. (a) Pattern centered at (x, y). (b) Pattern oriented with $\vec{\alpha}_x$. (c) Pattern scaled along the axes proportional to (s_x, s_y).

3.2.3. Training, Testing and Space Marginalization

In our discriminative learning approach parameter estimation involves searching a domain Ω for the most parameter values with the maximum posterior probability using a classifier D. Assuming an abstract model parametrization x and an input image I, the task can be formulated as:

$$\arg\max_x p(x|I) = \arg\max_x D(x, I), x \in \Omega \tag{3.11}$$

The framework is comprised of two phases: training and testing. The classifier D is constructed during an offline training phase. It is used to verify samples from the parameter search space and to find the optimal solution in the following testing phase.

Training: The objective of the training stage is to learn a classifier D from a given training set. This comprises out of pairs of labeled parameter instances $(x, y), y \in \{+1, -1\}$ and includes positive $(x, +1)$ and negative $(x, -1)$ samples. Positive and negative examples are obtained from a set of images associated with a ground truth annotation $(I, \overline{x})$. Positive samples are located in close proximity to the ground truth parameters, which is associated with a a distance measure d and threshold value δ_s for its deviation:

$$y = \begin{cases} 1 & if \quad d(x, \overline{x}) < \delta_s \\ -1 & otherwise \end{cases} \tag{3.12}$$

Weak learners h are employed to model the target distribution. Those are constructed from a pool of image features parametrized by the model:

$$h(x, f, \delta_f) = \begin{cases} 1 & if \quad f(x) < \delta_f \\ 0 & otherwise \end{cases} \tag{3.13}$$

Examples of features f relevant for our work were presented in section 3.2.2. The threshold δ_f are chosen during learning such that the minimum number of training samples are misclassified. Single weak learners h do not produce satisfactory classification results. Therefore, as illustrated in section 3.2.1, boosting algorithms are employed to select key weak learners and aggregate them to build a strong classifier D.

Testing: The objective of the testing stage is to identify the highest probable parameter values in a predefined search space domain Ω. The search domain is usually discretized by a set of hypotheses $\mathcal{H}$. The learned classifier D exhaustively computes the posterior probability for each sample (or hypothesis) and ranks each sample. The final result is either the top ranked sample or an aggregation of the top N samples. Exhaustive search provides robustness against local minima and assuming that the true solution is captured in $\mathcal{H}$, it also provides the optimal parameter estimation.

However, the exhaustive search strategy can become computationally expensive as its complexity increases exponentially with the dimensionality of the target parameters, which makes the estimation of high-dimensional parameters intangible. For instance, the problem of estimating the similarity transformation of an object in a three-dimensional Euclidean space, where the model has nine parameters $dim(x) = 9$ (3 for translation, 3 for rotation and 3 for scale) and if each dimension of the search domain Ω is discretized by only 10 values, the number of hypotheses to be tested is $|\mathcal{H}| = 10^9$.

To overcome this limitation, marginal space learning is employed, which drastically reduces the search domain while improving classification performance.

Marginal Space Learning: In order to avoid sampling the complete domain Ω we build upon the assumption that most of the posterior distribution over the parameter space is clustered in a small region of the high-dimensional space spanned by the parameters in x. This observation is exploited within the Marginal Space Learning (Marginal Space Learn-

ing (MSL)) framework [205], which breaks the original domain Ω into subsets of marginal spaces with increased dimensionality:

$$\Omega_1 \subset \Omega_2 \subset \cdots \subset \Omega_n = \Omega \tag{3.14}$$

where $dim(\Omega_1) \ll dim(\Omega)$ and $dim(\Omega_k) - dim(\Omega_{k-1})$ is small. A search in Ω_1 with a detector D_1 learned in this marginal space finds a subspace $\mathcal{C}_1 \subset \Omega_1$, which contains only the most probable parameter values and discards the rest of the space, such that $|\mathcal{C}_1| \ll |\mathcal{H}_1|$. Another stage of training and testing is performed in the extended $\mathcal{C}_1^e = \mathcal{C}_1 \times \mathcal{H}_2 \subset \Omega_2$ to obtain a restricted marginal space $\mathcal{C}_2 \subset \Omega_2$. The procedure ends when the final dimensionality of Ω is reached. In practice, the optimal arrangement for MSL sorts the marginal spaces in a descending order based on their variance. Learning parameters with low variance first will increase the overall precision of the detection.

3.3. Shape Regularization

Incorporating prior knowledge about the shape of interest is important to assure the segmentation result lies within the same domain as the training cases. In most segmentation algorithms the estimation can be divided in two main parts: data term and regularization. The data term deforms or updates the current model according to the current image. However this shape can be highly irregular and not resemble the model from the annotation database. To bring the shape into the domain of the annotations regularization is necessary. Many parametric and non-parametric methods have been developed to achieve this goal.

3.3.1. Statistical Shape Models

Statistical shape models were propose by Cook [33]. They represent linear models used to efficiently represent the variation of a specific shape both across the temporal domain and across patient population. In our case the shape is the point distribution model of the complete valvular apparatus.

The first stage involves aligning all shapes. This means the final model should capture the variations within the shapes by removing the Euclidean similarity transform between the shapes. As mentioned in chapter 3.1 our model consist of 11 surfaces. The first requirement is that all shapes have point correspondence and all shapes must contain the same number of shape points. An instance of A_k, with N_k vertices $\vec{p}^{A_k} \in \mathbb{R}^3$ is represented as $3 \times N_k$ element vector X:

$$X = (p_x^1, \ldots, p_x^{N_k}, p_y^1, \ldots, p_y^{N_k}, p_z^1, \ldots, p_z^{N_k})^T \tag{3.15}$$

As our shapes need to be considered invariant to Euclidean similarity transformations within a reference coordinate system to which all shapes are aligned we applied the General Procrustes Analysis (Generalized Procrustes Analysis (GPA)) [63] algorithm to achieve

this task. Although analytical solutions exists [83], the most popular approach to shape alignment is the GPA. The task is defined as optimization problem of each shape, relative to the mean:

$$D = \sum \parallel X_i - \overline{X} \parallel^2 \tag{3.16}$$

where, X_i is a shape vector from a training set of S samples and $\overline{X}$ the mean shape of this training set defined as:

$$\overline{X} = \frac{1}{S} \sum_{i=1}^{S} X_i \tag{3.17}$$

Two arbitrary shapes X_1 and X_2, with their center of gravity at the origin, are aligned with respect to the scale s and rotation R parameters by minimizing the following equation:

$$\arg\min_{s,R} D(s, R) = \arg\min_{s,R} \parallel T_{s,R}(X_1) - X_2 \parallel^2 \tag{3.18}$$

where the optimal solution is provided by the sum of squared differences between the points of X_2 and transformed points of X_1. The iterative solution to the GPA includes four steps:

1. Set an initial estimation of the mean shape, e.g. X_1
2. Align all shapes $X_1, \ldots, X_S$ with the current mean shape using Eq. 3.18
3. Re-estimate mean from aligned shapes with Eq. 3.17
4. If estimated mean has changed re-iterate from step 2.

Convergence is assumed if the current mean does not change significantly with respect to the previous estimation. Initially, all shapes $X_1, \ldots, X_S$ are translated into the origin and scaled such that $|X| = 1$, which will cause their alignment on a hypersphere after scale and rotation minimization. To avoid non-linearities and simplify the shape distribution description, the minimization is performed on the tangent space in respect to the mean. A simple approach to achieve this is to scale each shape with $\frac{1}{X_i \cdot \overline{X}}$ after step 2.

The initial shape space, which is the set of all shapes $X_1, \ldots, X_S$, spans a space of dimensionality $k \times n = 3 \times N_k$. After adjusting for translation, scale and rotation, which removes k, 1 and $1/2k(k-1)$ dimensions respectively, the dimensionality of the aligned subspace is equal to $kn - k - 1 - \frac{k(k-1)}{2}$ [160, 37].

Using the aligned training set of shapes $X_1, \ldots, X_S$, a model of their distribution enables generation of new plausible shapes and, vice-versa, examination of new shapes for plausibility. The goal is to model $p(X)$ using a linear model $X = M(\rho)$, where the parameter vector ρ lies in a lower dimensional space. An effective approach to dimensionality reduction is the Principal Component Analysis Principal Component Analysis (PCA). Assuming a Gaussian distribution of the population with the variance equal to:

$$\Sigma_X = \frac{1}{S}\sum_{i=1}^{S}(X_i - \overline{X})(X_i - \overline{X})^T \tag{3.19}$$

each shape can be approximated using the mean shape $\overline{X}$ and highest ranked eigenvectors ϕ_i, correspondent to largest eigenvalues λ_i of Σ_X:

$$X \approx \overline{X} + \Phi\rho \ , \ \ \rho = \Phi^T(X - \overline{X}) \tag{3.20}$$

where the vector ρ contains the model parameters, Φ is the matrix formed by the first t column Eigenvectors of the symmetric covariance matrix Σ_X. To enforce plausible shape generation with respect to the training set, ρ_i is bound by $\pm 3\sqrt{\lambda_i}$. The dimensionality of the model parameter vector ρ is equal to the number of Eigenvectors retained [171, 96, 131].

3.4. Patient-Specific Model Estimation

A hierarchical estimation approach is utilized to deduce model parameters, introduced in the previous section from 4D cardiac CT images. First, robust machine learning techniques are applied to estimate the global valves and anatomic landmarks parameters introduced in Eq. 3.1 and 3.2. Second, we present a novel anatomical constrained Multi-linear Shape Model (cMSM), which effectively captures the complex spatio-temporal variation of all valves. Finally, the cMSM is applied in a learning-based framework to estimate the complete valve model described in Eq. 3.3.

3.4.1. Global Motion estimation

The global motion estimation is formulated as a classification problem in order to estimate $\boldsymbol{B}_t$ for each time step t independently from the corresponding volumes $I(t)$. The probability $p(\boldsymbol{B}(t)|I(t))$ can be modeled by a learned detector D, which evaluates and scores a large number of hypotheses for $\boldsymbol{B}_t$. To avoid an exhaustive search along a nine-dimensional space of $\boldsymbol{B}_t$ we apply the Marginal Space Learning framework [205] and decompose the original parameter space into a subset of increasing marginal spaces:

$$\Omega_1 \subset \Omega_2 \subset \cdots \subset \Omega_n = \Omega \tag{3.21}$$

The nine-dimensional space described by the similarity transform in a three-dimensional Euclidean space is decomposed as follows:

$$\begin{aligned} \Sigma_1 &= (c_x, c_y, c_z) \\ \Sigma_2 &= (c_x, c_y, c_z, \vec{\alpha}_x, \vec{\alpha}_y, \vec{\alpha}_z) \\ \Sigma_3 &= (c_x, c_y, c_z, \vec{\alpha}_x, \vec{\alpha}_y, \vec{\alpha}_z, s_x, s_y, s_z) \end{aligned} \tag{3.22}$$

where Σ_1 represents the position marginal space, Σ_2 the position + orientation marginal

space and Σ_3 the position + orientation + scale marginal space, which coincides with the original domain. Detectors are trained using the Probabilistic Boosting Tree using Haar-like and Steerable Features for each marginal space D_1, D_2 and D_3, and $\boldsymbol{B}_t$ is estimated by gradually increasing the dimensionality. As described in [205], the 100 highest scored candidates are retained in Σ_1, 50 in Σ_2 and 25 in Σ_3, such that the smallest subgroup which is likely to include the optimal solution is preserved.

To obtain a temporally consistent global location a RANSAC estimator is employed. To suppress temporally inconsistencies, we assume a constant model for the cardiac motion, which drives the global movement of the entire valvular apparatus. From randomly sampled candidates, the one yielding the maximum number of inliers is picked as the final motion. Inliers are considered within a distance of $\sigma = 7mm$ from the current candidate and extracted at each time step t. The procedure is applied for each valve separately, in order to obtain the resulting time-coherent similarity transform $\boldsymbol{B}_t$ assuming small displacements between consecutive frames.

3.4.2. Landmark Location and Motion Estimation

The landmarks parameters are estimated within the Marginal Space Learning framework [205] using an algorithm called Trajectory Spectrum Learning (TSL), similar to [87]. Hereby the landmark motions are represented in the frequency domain instead of the Euclidean space. Therefore the motion estimation problem is formulated as spectrum learning and detection in the trajectory space. The object localization and motion estimation, referred traditionally as detection and tracking are solved simultaneously.

The trajectory $\boldsymbol{L}_n(\boldsymbol{B})$ of each landmark can be uniquely represented by the concatenation of its discrete Fourier transform (DFT) coefficients

$$\vec{s^j} = [\vec{s^j}(0), \vec{s^j}(1), \cdots, \vec{s^j}(n-1)] \tag{3.23}$$

obtained through the DFT equation:

$$\vec{s^j}(f) = \sum_{t=0}^{n-1} \boldsymbol{L}_n(\boldsymbol{B})(t) e^{\frac{-j2\pi tf}{n}} \tag{3.24}$$

where $\vec{s^j}(f) \in \mathcal{C}^3$ is the frequency spectrum of the x, y, and z components of the trajectory $\boldsymbol{L}_n(\boldsymbol{B})$, and $f = 0, 1, \cdots, n-1$. A trajectory $\boldsymbol{L}_n(\boldsymbol{B})$ can be exactly reconstructed from the spectral coefficients $\vec{s^j}$ applying the inverse DFT:

$$\boldsymbol{L}_n(\boldsymbol{B}) = \sum_{f=0}^{n-1} \vec{s^j}(f) e^{\frac{j2\pi tf}{n}} \tag{3.25}$$

By decomposing the full trajectory space into orthogonal subspaces defined by generic bases, such as the Discrete Fourier Transform (DFT), the obtained representation is shown to be compact especially for periodic motions, such as the movements of the heart valves.

This resulting compact representation allows efficient learning and optimization in its marginal spaces. In the training stage, local features are extended in the temporal domain to integrate the time coherence constraint. Thereby simple gradient and intensity information is extracted from the image forming three-dimensional features F^{3D}. As the motion of the landmarks is assumed to be locally coherent, F^{3D} is applied in a temporal neighborhood $t - T$ to $t + T$. The final value of the Local-Spatial-Temporal (LST) feature is the result of time integration using a set of linear kernels τ, which weight the spatial features F^{3D} according to their distance from the current frame t.

As described earlier the landmark trajectory is represented in the frequency space. Due to their periodic motion a small set of the frequency components is sufficient to represent their motion. These frequency subspaces $\Sigma^{(k)}$ are efficiently represented by a set of corresponding hypotheses $\mathcal{H}^{(k)}$ obtained from the training set. The pruned search space, restricted to ζ frequency components, enables efficient learning and optimization:

$$\Omega_{r-1} = \mathcal{H}^{(0)} \times \mathcal{H}^{(1)} \times \ldots \times \mathcal{H}^{(r-1)}, r = |\zeta| \tag{3.26}$$

The training algorithm starts by learning the posterior probability distribution in the marginal space Ω_0. Subsequently, the learned detector D_0 is applied to identify high probable candidates $\mathcal{C}_0$ from the hypotheses $\mathcal{H}^{(0)}$. In the following step, the dimensionality of the space is increased by adding the next spectrum component. For each marginal space Σ_k, corresponding discriminative classifiers D_k are trained on sets of positives and negatives. The Local-Spatial-Temporal (LST) features are selected via the Probabilistic Boosting Tree to form the strong classifier D_k.

In order to estimate the final trajectory of a landmark we start from the zero-spectrum and incrementally estimate the magnitude and phase of each frequency component $\vec{s}(k)$. At stage k, the corresponding robust classifier D_k is exhaustively scanned over the potential candidates $\mathcal{C}_{k-1} \times \mathcal{H}^{(k)}$. The final trajectory is reported as the average of all elements in $\mathcal{C}_{r-1}$.

3.4.3. Constrained Multi-linear Shape Model

Multilinear modeling enables the decomposition of a shape space in a temporal and spatial component in contrast to active shape models (ASM) where both are coupled. In this section we present a Multi-linear Multi-linear Principal Component Analysis (MPCA) (Multilinear Principle Component Analysis) and Multi-linear Independent Component Analysis (MICA) (Multi-linear Independent Component Analysis) shape model of all valves which is conditioned by anatomical measurements.

Shape Space. In order to construct the shape model, all shapes $\boldsymbol{V}$ are aligned by calculating the mean sequence model and aligning them using General Procrustes Analysis (GPA). This transform is utilized to align all shapes in the sequence. The normalized shapes are represented as third-order tensors $\boldsymbol{D} \in \mathbb{R}^{(\mathrm{S}\times\mathrm{T}\times\mathrm{P})}$, where S is the number of patients, T is

the frame number inside a multi phase sequence and P represents the number of shape points. The final third-order tensors $\boldsymbol{D}$ is constructed as follows:

$$\mathcal{D} = \mathcal{Z} \times_1 \mathbf{U}_{\text{patient}} \times_2 \mathbf{U}_{\text{motion}} \times_3 \mathbf{U}_{\text{points}} \tag{3.27}$$

where $\mathbf{U}_{\text{patient}}$ is representing the patient modes, $\mathbf{U}_{\text{motion}}$ the motion modes, $\mathbf{U}_{\text{points}}$ the points modes and $\mathcal{Z}$ the core tensor. As mentioned by Zhu et al. [210] the motion subspace due to its non-Gaussian distribution is decomposed using ICA and the patient and points space using PCA. We use the fixed point algorithm to perform the Independent Component Analysis ([85]). Thereby the equation 3.27 is modified by introducing the linear static transformation $\mathbf{W}$.

$$\begin{aligned} \mathcal{D} &= \mathcal{Z} \times_1 \mathbf{U}_{\text{patient}} \times_2 \mathbf{U}_{\text{motion}} \mathbf{W}^{-1} \mathbf{W} \times_3 \mathbf{U}_{\text{points}} \\ &= (\mathcal{Z} \times_2 \mathbf{W}) \times_1 \mathbf{U}_{\text{patient}} \times_2 \mathbf{U}_{\text{motion}} \mathbf{W}^{-1} \times_3 \mathbf{U}_{\text{points}} \\ &= \mathcal{S} \times_1 \mathbf{U}_{\text{patient}} \times_2 \mathbf{C}_{\text{motion}} \times_3 \mathbf{U}_{\text{points}} \end{aligned} \tag{3.28}$$

Constrained Model Estimation. A crucial step in our hierarchical model estimation algorithm is to advance from one model hierarchy layer to the next finer. This step is especially important when moving from the anatomical landmark representation to the dense surface mesh models. Instead of using a warping technique, like the thin-plate spline interpolation, to map a mean mesh model to the location of the landmarks we use a Bayesian approach to estimate the dense surface meshes from meaningful clinical measures. A set of anatomical measurements $\boldsymbol{M}$ $(m_1, m_2, \ldots, m_R)$ extracted from the non-linear valve model used to condition a surface parametrization $\boldsymbol{V}_q$ $(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_K)$ [13]. In the context of the aortic valve root $\boldsymbol{V}_1$ three measurements are used: 1) inter-commissure distance, 2) hinge-leaflet tip distance and 3) inter-hinges distance (see figure 3.18).

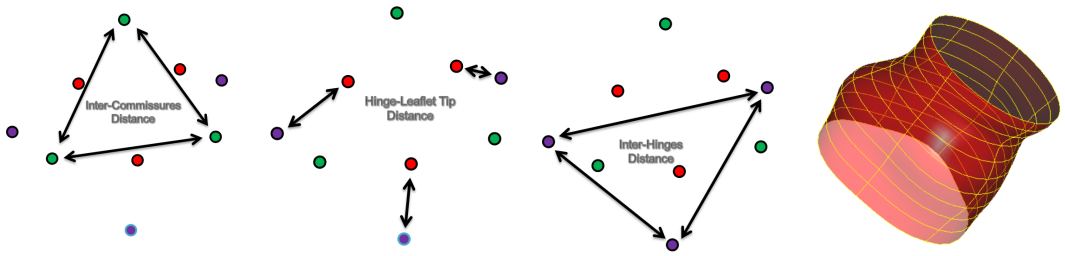


Figure 3.18.: Anatomical measurements extracted from the aortic valve anatomical landmarks model (a) inter-commissures distance, b) hinge-leaflet tip distance and c) inter-hinge distance) in order to constrain the full surface model d). The green points are representing the aortic valve commissures, the purple point the hinges and the red the leaflet tips.

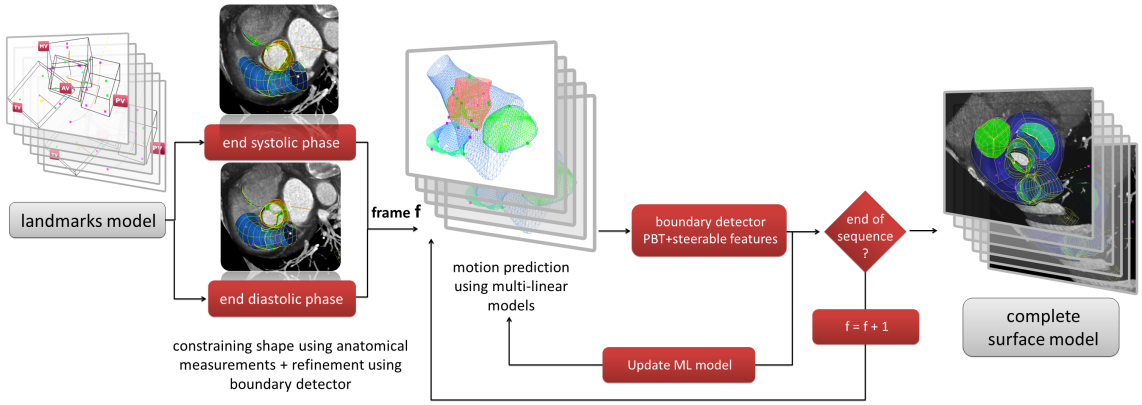


Figure 3.19.: Diagram depicting the estimation process of the complete valve model during a full cardiac cycle.

Assuming the joint multivariate distribution $(\boldsymbol{V}_q \,|\boldsymbol{M})$ follows a Gaussian distribution a conditioned surface $\boldsymbol{V}_q^M$, containing the anatomical measurements $\boldsymbol{M}$, can be estimated as follows:

$$\boldsymbol{V}_q^M = \mu_{\boldsymbol{V}_q} + \boldsymbol{\Sigma}_{\boldsymbol{V}_q\boldsymbol{M}}\, \boldsymbol{\Sigma}_{\boldsymbol{M}\boldsymbol{M}}^{-1} (\boldsymbol{M} - \mu_{\boldsymbol{M}}) \tag{3.29}$$

where $\mu_{\boldsymbol{V}_q}$ is the mean surface parameterization from all training sets of the valve surface $\boldsymbol{V}_q$, $\mu_{\boldsymbol{M}}$ the mean of the measurements $\boldsymbol{M}$ in the training set, $\boldsymbol{\Sigma}_{\boldsymbol{V}_q\boldsymbol{M}}$ the covariance matrix between $\boldsymbol{V}_q$ and $\boldsymbol{M}$. The constrained surface $\boldsymbol{V}_q^M$ is used to reconstruct the dynamic motion surface model of the whole sequence. Therefore we first estimate the patient modes $\mathbf{u}_{\text{patient}}$ and then use them to reconstruct $\boldsymbol{V}_q\,(\boldsymbol{L},\boldsymbol{B})$.

$$\mathbf{u}_{\text{patient}} = \boldsymbol{V}_q^M\ \mathbf{T}_{(1)}^{-1} \qquad \mathcal{T} = \mathcal{S} \times_2 \mathbf{C}_{\text{motion}} \times_3 \mathbf{U}_{\text{points}} \tag{3.30}$$

where $\mathbf{T}_{(1)}^{-1}$ is the pseudo-inverse of the tensor $\mathcal{T}$ flattened along Z $Z \in 1,\ldots,T$ modes and $\mathbf{C}_{\text{motion}}$ the Z dimensional motion modes. The complete surface model for the complete sequence can be extracted by a tensor multiplication:

$$\boldsymbol{V}_q\,(\boldsymbol{L},\boldsymbol{B}) = \mathcal{S} \times_1 \mathbf{u}_{\text{patient}} \times_2 \mathbf{C}_{\text{motion}} \times_3 \mathbf{U}_{\text{points}} \tag{3.31}$$

3.4.4. Complete Valve Model Estimation

The final stage in our hierarchical model estimation algorithm is the estimation of the complete surface model $\boldsymbol{V}_q\,(\boldsymbol{L},\boldsymbol{B})$. The shape model of each valve is first initialized in the End-Diastole (ED) and End-Systole (ES) phases of the cardiac cycle using anatomical measurements $\boldsymbol{M}$ defined between the landmarks $\boldsymbol{L}_1 \cdots \boldsymbol{L}_{33}$. In the case of the aortic valve the shape is conditioned using three anatomical measurements extracted from the previously estimated landmark model: $\boldsymbol{M} = \{m_1, m_2, m_3\}$($m_1$-inter-commissure distance, m_2-hinge-

commissure plane distance, m_3-hinge-commissure plane angle). The initialized model is refined using a boundary detector D learned using the probabilistic boosting-tree [179] and steerable features [50]. The detector D evaluates hypotheses for each discrete boundary point along its corresponding normal direction. The new boundary points are set to the hypotheses with maximal probability. To guarantee physiologically compliant results, the final model for each frame is obtained after projecting the estimated points to the multi-linear shape space described in section 3.4.3. Thereby the multi-linear shape space is used as a parametric space limiting the variability of the final shape model. Starting from the estimation results in the ED and ES phases, model parameters can be predicted in the remaining frames by utilizing the multi-linear shape model as described in section 3.4.3. Thus an initialization of the models is available in the remaining frames of the sequence and it is conditioned on the estimation results in the ED and ES frames. Starting from the neighboring frames $t_{ED+1}, t_{ED-1}, t_{ES+1}, t_{ES-1}$ the initialization is refined using the boundary detector D and the result projected to the parametric multi-linear shape space. Thereby the patient specific modes $\mathbf{u}_{\text{patient}}$ are updated and thus the predictions in the remaining frames are more accurate as the variability or the dynamic shape model was reduced. The procedure is repeated until the full 4D model is estimated for the complete sequence (see Figure 3.19).

3.5. Experimental Results

The accuracy of the proposed method is evaluated using cardiac CT data sets from patients affected by a large spectrum of cardiovascular and valvular heart diseases. Among the included pathologies are: regurgitation, stenosis, prolapse and aortic root dilation. The ECG gated cardiac CT sequences included multiple volumes per cardiac cycle, where each volume contains 80-350 slices with 153×153 to 512×512 pixels. The in-slice resolution is isotropic and varies between 0.28 to 1.00mm with a slice thickness from 0.4 to 2.0mm. The imaging data set includes 64 cardiac CT studies (640 volumes) which were collected from several medical centers around the world. Using heterogeneous imaging protocols, cardiac CT exams were performed with Siemens Somatom Sensation or Definition scanners. Each sequence was acquired over one cardiac cycle and consisted of ten volumes. Only data sets which contained a contrast agent and all valves were visible were used. In order to keep the radiation dose low during the acquisition most of the data sets had one peak dose at either the ED or ES phase and a low dose during the rest of the cardiac cycle. Therefore the best visibility of the valves was during the peak dose phase and a moderate quality during the remaining cycle. The ground-truth for training and testing was obtained through an incremental annotation process. Therefore, each volume in our data set is associated with an annotation obtained through an expert-guided process that includes the following steps:

- the anatomical landmark motion model is manually determined by placing each anatomical landmark (see section 3.1.2) at the correct location in the entire cardiac

cycle of a given study. From the annotated anatomical landmark model, the global dynamic motion model $\boldsymbol{B}_t$ is determined as described in section 3.1.2.

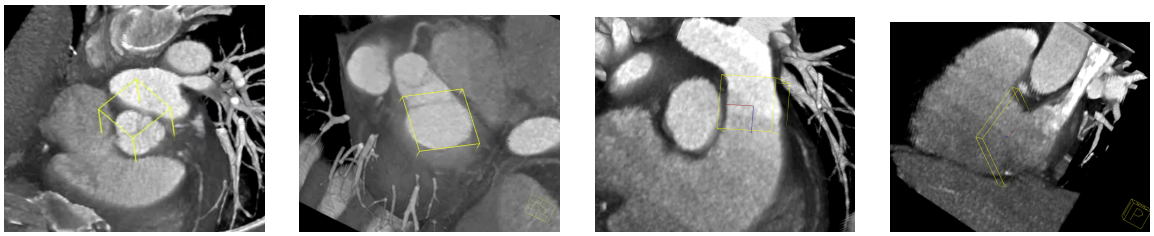
- the complete valve model is initialized through its mean model placed at the correct image location, expressed by the thin-plate-spline transform estimated from the previously annotated anatomical landmark model (see section 3.1.3).

- the annotation of the complete valve model is manually adjusted to delineate the true valves boundaries over the entire cardiac cycle (see section 3.4.4). Complex resampling algorithms specialized for each valve were developed to ensure temporal and spatial consistency during the annotation process.

In addition each evaluation is done using three-fold cross validation.

An inter-user experiment was conducted on a randomly selected subset of sixteen studies for the aortic and mitral valve. The patient-specific landmark valve models $\boldsymbol{L}_n(\boldsymbol{B})$ were manually fitted by four experienced users. The ground-truth was assumed to be the mean of the four user annotations. A landmark error of $1.53mm \pm 0.93$ for the aortic valve and $1.97mm \pm 1.4$ for the mitral valve was observed.

The performance of the global dynamic motion estimation, $\boldsymbol{B}_t$, described in section 3.4.1, is evaluated in two distinct experiments. First, the overall detection precision is quantified at the box corners of the detected time-dependent similarity transformation. The average Euclidean distance between the eight bounding box points, defined by the similarity transform parameters $\{(c_x, c_y, c_z)_i, (\vec{\alpha}_x, \vec{\alpha}_y, \vec{\alpha}_z)_i, (s_x, s_y, s_z)_i\}$ and the ground-truth box is reported. Table 3.1 illustrates the mean errors and corresponding standard deviations distributed over the four valves. Examples of estimation results are given in Figure 3.20.



(a) (b) (c) (d)

Figure 3.20.: Examples of global dynamic motion estimation in cardiac CT: (a) aortic valve, (b) mitral valve, (c) pulmonary valve, (d) tricuspid valve.

In a second experiment, the accuracy of the individual detection stages is investigated. Absolute differences between estimated and ground-truth parameters of the position, orientation, and scale are reported in table 3.2. The 80% column represents the 80th percentile of the error values. Please note that in order to speed up the algorithm, the estimation of the global location and rigid motion is always performed on downsampled data with an isotropic resolution of 3mm.

	Aortic Valve	Mitral Valve	Pulmonary Valve	Tricuspid Valve
mean error [mm]	4.32	6.72	7.72	8.12
STD [mm]	1.90	2.21	2.3	3.2

Table 3.1.: Accuracy of the global location and rigid motion estimation, quantified from the box corners and reported using the mean error and standard deviation distribution over each valve.

	Mean / STD	Median	80%
Position [mm]	3.09±3.02	2.33	3.23
Orientation [deg]	9.72±5.98	7.93	10.73
Scale [mm]	6.50±4.19	5.09	7.81

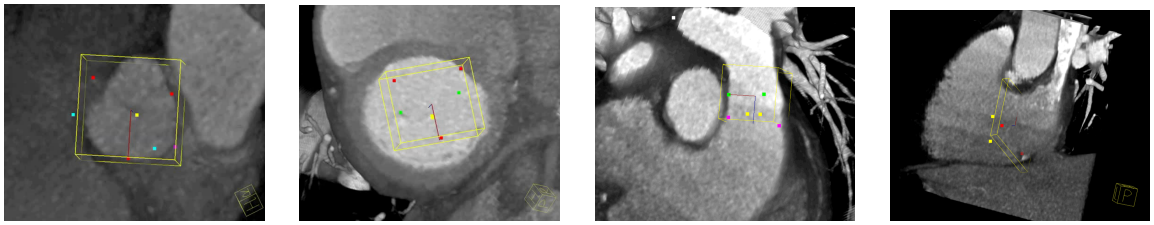
Table 3.2.: Accuracy of the global location and rigid motion estimation reported separately for position, orientation and scale.

The accuracy of the anatomical landmark motion model, $\boldsymbol{L}_n(\boldsymbol{B})$, presented in section 3.1.2 is measured using the Euclidean distance between detected and corresponding ground-truth landmark trajectories. Table 3.3 demonstrates the precision expressed in mean errors and standard deviations, distributed over the four valves. Note that reported values are obtained by averaging the performance of individual landmarks with respect to the corresponding valve. Examples of estimation results are given in Figure 3.21. The detection was performed on volumes resampled to an isotropic resolution of 1.00mm. Thus our automated landmark estimation error (1.53mm for the aortic valve and 1.97mm for the mitral valve) is slightly above (1.13mm for the aortic valve and 0.78mm for the mitral valve) the intra-user variability error (1.53mm for the aortic valve and 1.97mm for the mitral valve).

	Aortic Valve	Mitral Valve	Pulmonary Valve	Tricuspid Valve
mean error [mm]	2.65	2.75	3.50	3.59
STD [mm]	1.50	1.19	2.53	2.55

Table 3.3.: Accuracy of the non-rigid landmark motion estimation, quantified by the Euclidean distance and reported using the mean error and standard deviation distribution over each valve.

The accuracy of the algorithm in section 3.1.3 to estimate the comprehensive valvular model, $\boldsymbol{V}_q(\boldsymbol{L}, \boldsymbol{B})$, (see section 3.4.4) is evaluated by utilizing the point-to-mesh distance. For each point on a surface $\boldsymbol{V}_q$, we search for the closest point on the other surface to calculate the Euclidean distance. To guarantee a symmetric measurement, the point-to-mesh distance is calculated in two directions, from detected to ground-truth surfaces and vice versa. Table 3.4 contains the mean error and standard deviation distributed over the



(a) (b) (c) (d)

Figure 3.21.: Examples of the anatomical landmark motion estimation in cardiac CT: (a) aortic valve, (b) mitral valve, (c) pulmonary valve, (d) tricuspid valve. The colored points are showing the landmarks for each valve.

four valves. The detection of the comprehensive valves model was performed on volumes resampled to an isotropic resolution of $1mm$. Examples of estimation results are given in Figure 3.22.

	Aortic Valve	Mitral Valve	Pulmonary Valve	Tricuspid Valve
mean error [mm]	1.22	1.32	1.35	1.40
STD [mm]	0.38	0.57	0.9	1.41

Table 3.4.: Accuracy of the comprehensive valve model estimation, quantified by the point-to-mesh distance and reported using the mean error and standard deviation distribution over each valve.

In the second experiment, we compare our new shape estimation approach with two other methods. Thereby the error is measured as the point-to-mesh distance between the estimated and ground-truth mesh. For all methods the estimation of the dynamic global motion $\boldsymbol{B}_t$ and the anatomical landmark model $\boldsymbol{L}_n(\boldsymbol{B})$ is done as described in chapter 3.1.2. The results, shown in Table 3.9 corroborate that our constrained ML PCA-ICA shape estimation approach achieves best performance, compared to a regular ML PCA-Independent Component Analysis (ICA) method and a standard frame-wise estimation procedure (tracking by detection). Within three minutes a complete personalized dynamic model of all valves is estimated with an average accuracy of 1.24mm. The full valvular model together with the four chambers of the heart is illustrated in Figure 3.22 and 3.41.

Important clinical parameters are extracted from the personalized model in the right heart. They include right-ventricle outflow tract (RVOT) radius, bifurcation radius, tricuspid valve area and a joint measurement of the two valves, the pulmonary and tricuspid valve distance. Quantitative comparison is shown in figure 3.24 by comparing ground truth measurements and the estimated, demonstrating a strong correlation.

Finally we show quantitative comparison between a patient suffering from aortic valve regurgitation, a healthy patient and a post-operative patient who underwent a Ross operation. An important clinical measurement, the valvular area, extracted from the personal-

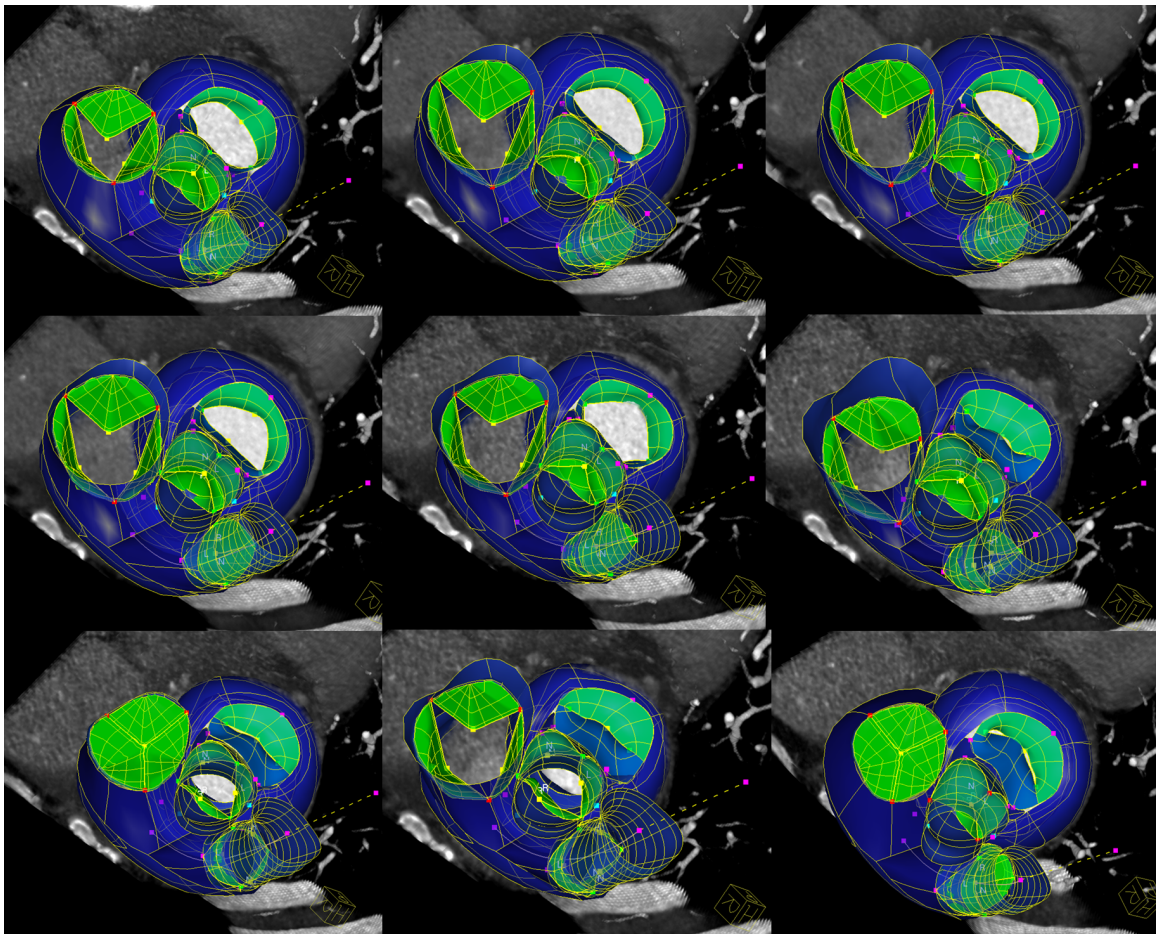


Figure 3.22.: Examples of the complete valves model estimation in cardiac CT of all heart valves during one cardiac sequence.

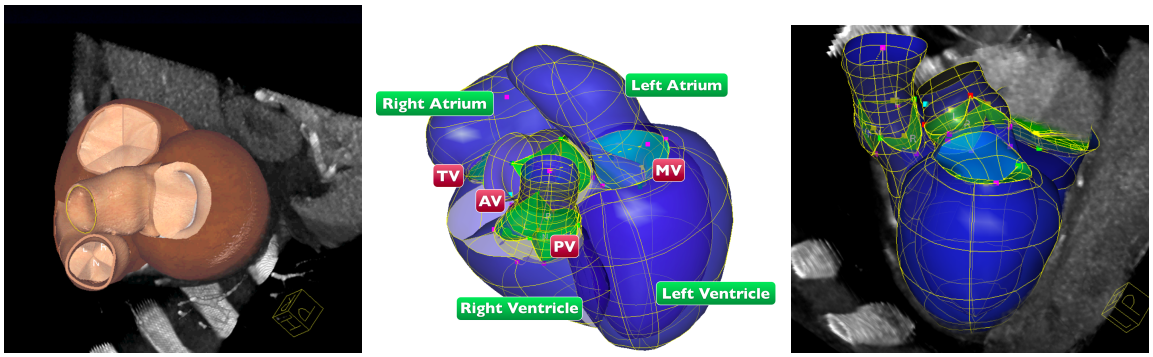


Figure 3.23.: Examples of estimated personalized model from a multiphase CT sequence. The images are extracted from the end-systolic phase.

	Mean	STD	Median
Tracking by Detection [mm]	1.52	0.98	1.47
ML PCA-ICA [mm]	1.39	0.91	1.32
cML PCA-ICA [mm]	1.24	0.91	1.18

Table 3.5.: System precision for valve model estimation averaged over all valves for comprehensive surface assessment.

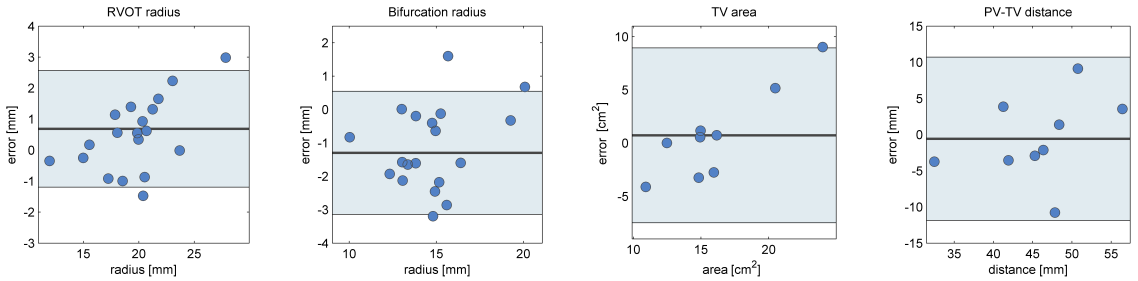


Figure 3.24.: Bland Altman plots for a) right ventricle output tract diameter, b) pulmonary valve bifurcation diameter, c) tricuspid valve area and d) distance between pulmonary and tricuspid valve. The ground truth measurements, derived from the models annotated by clinical experts, were compared with measurements derived from our automatically estimated models.

ized aortic and pulmonary valve model, demonstrated in figure 3.25, confirms a successful outcome since no regurgitation is observed at the aortic valve.

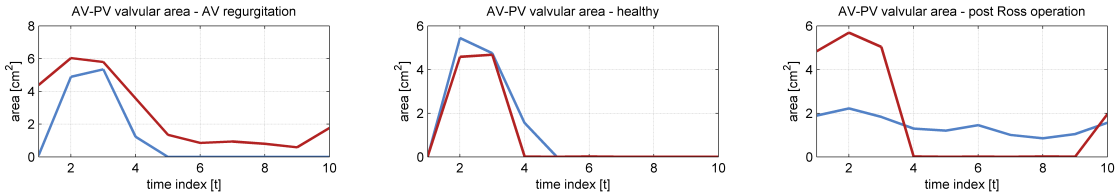


Figure 3.25.: Measurements of aortic (AV) and pulmonary valve (PV) area obtained from a patient with aortic valve regurgitaion (left), a healthy patient (middle) and a post Ross operation patient (right). The red graph is representing the aortic valve and the blue the pulmonary.

3.6. Tissue Characterization and Volumetric Model Estimation

For TAVI the interaction of calcification and leaflet tissue is an important predictor of post-operative success [111]. Thus having the capability to delineate tissues within the aortic valve from pre-operative data will advance the current clinical workflow for TAVI planning. In this chapter we propose an extension of the modeling framework presented in

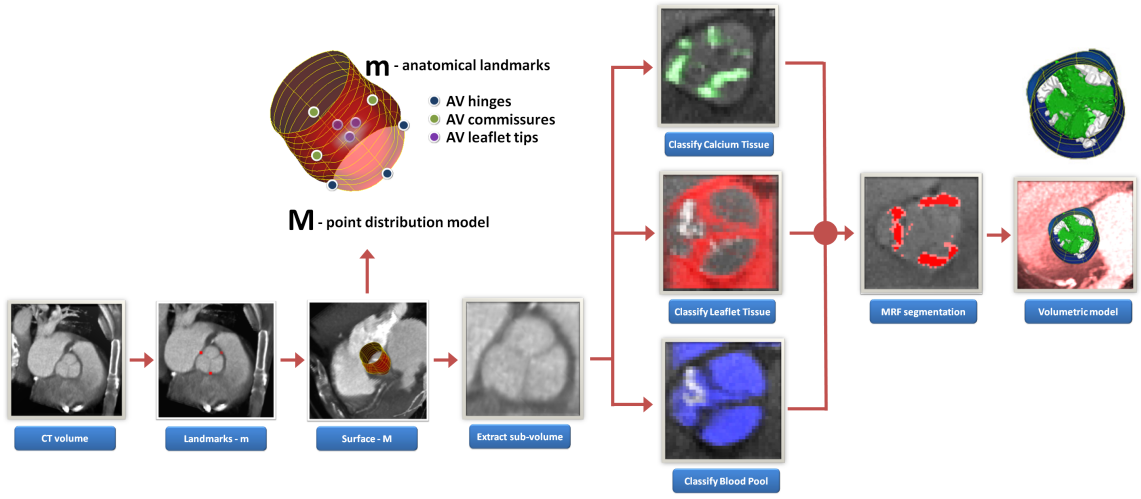


Figure 3.26.: Diagram showing the model estimation approach for the volumetric aortic valve model consisting of 9 anatomical landmarks m (3 commissures, 3 hinges and 3 leaflet tips), the aortic root surface M and the final volumetric aortic leaflet models.

3.1. First we apply the current framework of the physiological surface model of the aortic valve to a patient population of heavily stenotic AV patients. Afterward we propose a novel segmentation method based on discriminative learning-based methods to delineate the three tissue types within the aortic valve: blood pool, calcification and leaflet tissue. Based on our segmentation several clinical applications can be developed. Especially the association of calcium within the valve have been shown to have significant impact on the two main complications associated with TAVI: para-valvular leakages and stroke [73, 111].

3.6.1. Multi-class tissue Classification

We formulate our problem as a 3-class classification. Within the aortic valve root M we compute custom features x_i for each sample voxel i (see section 3.6.1) and assign the sample to one of three classes: calcification C_C, leaflet tissue C_L and blood pool C_B. We aim to finding a learning model $\mathcal{H}$ such that

$$\mathcal{H}(x_i) = y_i \qquad y_i \in \{C_C, C_L, C_B\} \tag{3.32}$$

where y_i is the class label for voxel i. We utilize binary boosting classifiers using geometric and appearance features to discriminate the voxels. In order to obtain the final class label, 1-vs-all approach is utilized. Herby the voxel i is assigned to the class with the maximum probability response of the classifier $\mathcal{H}$.

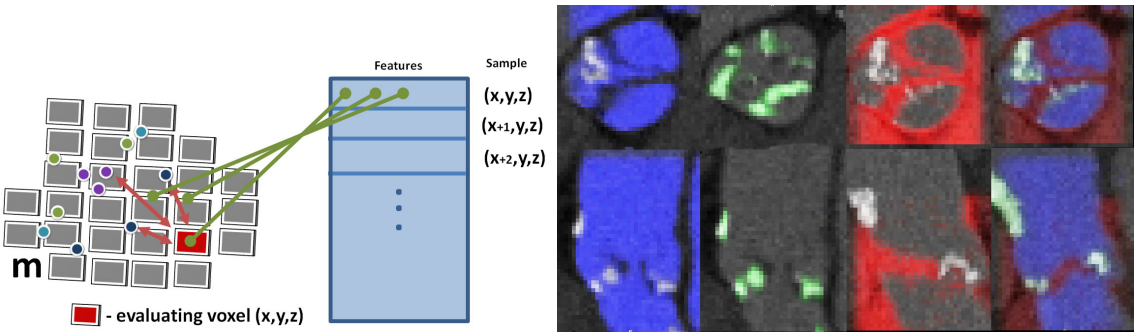


Figure 3.27.: Left: Custom geometric and data features utilized to classify tissues within the aortic valve root. Right: Classifier responses for different classes of tissues overlayed on the extracted subvolume. The probability map for the blood pool is overlayed on the blue channel, leaflet tissue in the red and calcium in the green color channel.

Geometric and Data Features

We use geometric and data driven features $x_i \in \mathbb{R}^{10}$ to discriminate the samples. For each voxel distances to the previously estimated landmarks m are used as geometric features and the image intensity at the current voxel position as appearance feature (see figure 3.26).

$$x_i(0) = I(i) \qquad x_i(j = 1 \dots 9) = ||i - m_j|| \tag{3.33}$$

Training

In order to train the binary classifiers $\mathcal{C}_C$, $\mathcal{C}_L$ and $\mathcal{C}_B$, each positive example x_i is assigned with a class label $y_i \in \{C_C, C_L, C_B\}$ extracted from user annotations. The calcification examples are generated by an expert user defining a volume specific threshold. Only responses within the previously estimated aortic valve root model M were taken. The ground-truth for the leaflet tissue position is extracted using a semi-automatic segmentation approach. The blood pool is assumed to be the remaining voxels when calcification and leaflet tissue is subtracted from the whole set of voxels within the aortic valve root M. Thus we train three binary classifiers: $\mathcal{C}_C$ (calcium), $\mathcal{C}_L$ (leaflets) and $\mathcal{C}_B$ (blood pool) using the probabilistic boosting tree [178]. The feature responses for each class C_C, C_L and C_B are shown in Figure 3.26.

Testing

First landmarks m and the point distribution model M are estimated. All pixels inside the aortic valve are evaluated using the classifiers $\mathcal{C}_C$, $\mathcal{C}_B$ and $\mathcal{C}_L$. The final voxel i is assigned to the class label with maximum probability.

3.6.2. Volumetric Leaflet tissue segmentation

The final volumetric segmentation of the leaflet tissue is formulated as a Markov random field (MRF) and optimized using graph-cuts [18].

Graph cuts in image segmentation

Graph cuts were introduced by Boykov et al. [18, 19, 17, 20] and utilized to efficiently solve many tasks in computer vision applications such as segmentation, image denoising and stereo reconstruction. These problems can be formulated as an energy minimization task and further reduced to instances of the maximum flow problem in a graph [17]. Using the max-flow min-cut theorem [6] the solution can be found efficiently and corresponds to the maximum a posterior of the estimate [101, 100]. Binary problems such as a segmentation task (0-background and 1-foreground) can be computed exactly.

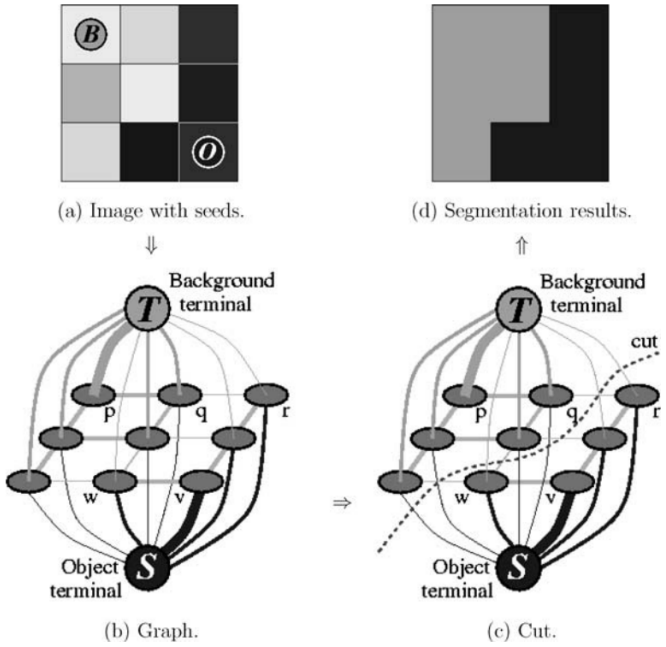


Figure 3.28.: A simple 2D segmentation example using graph-cuts for a 3×3 image from Boykov [20]. The seeds are $\boldsymbol{O} = v$ and $\boldsymbol{B} = p$. The cost of each edge is reflected by the thickness of the edge. The regional term and hard constraints define the costs of t-links. The boundary term defines the costs of n-links. Inexpensive edges are attractive choices for the minimum cost cut. A globally optimal segmentation satisfying hard constraints can be computed efficiently in low-order polynomial time using max-flow/min-cut algorithms on graphs ([49, 62, 31]

Let $A = (A_1, \ldots, A_p, \ldots, A_P)$ be a binary vector whose components A_p specify binary assignments to image pixels $p \in P$. Each A_p can be either object (1) or background (0).

Thus vector A defines a segmentation. The soft constraints that we impose on boundary and region properties of A are described by the cost function $E(A)$:

$$\boldsymbol{E}(A) = \lambda\, \boldsymbol{R}(A) + \boldsymbol{B}(A) \tag{3.34}$$

where

$$\begin{aligned} \boldsymbol{R}(A) &= \sum_{p \in P} R_p(A_p) \\ \boldsymbol{B}(A) &= \sum_{p,q \in \boldsymbol{N}} B_{p,q}\, \delta(A_p, A_q) \end{aligned}$$

and

$$\delta(A_p, A_q) = \begin{cases} 1 & if A_p \neq A_q \\ 0 & otherwise \end{cases} \tag{3.35}$$

The coefficient $\lambda \leq 0$ adjusts the relative importance of the region properties term $R(A)$ versus the boundary properties term $B(A)$. The term $R(A)$ assumes that the individual penalties for assigning pixel p as an object and background, $R_p(object)$ and $R_p(background)$, are given. E.g. $R_p(background)$ may suggest how a property of the pixel (e.g. intensity value at its location or some other form of posterior probability of belonging to the background) fits into the overall object and background model (e.g. in the case of intensities the histogram can be used). The term $B(A)$ comprises the boundary properties of segmentation A. Coefficient $B_{p,q} = 0$ should be interpreted as a penalty for a discontinuity between p and q. Usually $B_{p,q}$ is large when pixels p and q are similar (e.g. in their intensity) and $B_{p,q}$ is close to zero when the two are dissimilar. The cost of $B_{p,q}$ can also decrease as a function of distance between p and q. Costs $B_{p,q}$ may be based on local intensity gradient, gradient direction, and other criteria. Using the min-max flow algorithm [49, 62] a globally optimal binary segmentation can be computed by minimizing the energy in respect to the hard constraints.

In the context of segmentation using graph-cuts an undirected graph $G = \langle V, E \rangle$ is defined by a set of nodes (V) and a set of undirected edges (E) that connect these nodes. An example of a graph that we use in this section is shown in Figure 3.28. For each edge $e \in E$ the graph is assigned a non negative weight (cost) w_e. There are also two special nodes called terminals. A cut is a subset of edges $C \subset E$ such that the terminals become separated on the induced graph $G(C) = \langle V, E \setminus C \rangle$. It is normal in combinatorial optimization to define the cost of a cut as the sum of the costs of the edges that it severs

$$\|C\| = \sum_{c \in \boldsymbol{C}} w_e \tag{3.36}$$

Graph cut formalism is well suited for segmentation of images. In fact, it is completely appropriate for 3 dimensional volumes. The nodes of the graph can represent voxels and the edges can represent any neighborhood relationship between the pixels. A cut parti-

tions the nodes in the graph. As illustrated in Figure 3.28, this partitioning corresponds to a segmentation of an underlying image or volume. A minimum cost cut generates a segmentation that is optimal in terms of properties that are built into the edge weights. The technique is based on a well-known combinatorial optimization fact that a globally minimum cut of a graph with two terminals can be computed efficiently in low-order polynomial time [18, 19, 17, 20].

Leaflet segmentation

In the context of the volumetric aortic valve leaflet delineation the segmentation problem can be viewed as a labeling process to label the voxel set Q by minimizing an energy function:

$$E(L) = \sum_{p \in Q} D_p(f_p) + \sum_{q \in N(p)} V_{p,q}(f_p, f_q) \tag{3.37}$$

where $E(L)$ is the energy, p and q are voxels, N is the neighborhood formed from the vertex connectivity, $D_p(f_p)$ measures the cost of assigning the label f_p to pixel p, and $V_{p,q}$ measures the cost of assigning the labels f_p, f_q to the adjacent pixels p, q.

The positive and negative seeds are used from the tissue classification stage. Voxels in the regions outside the estimated aortic root model M are set as negative seeds. Voxels classified as leaflet voxels and as calcified regions are set as positives. The binary solutions assigns the uncertain voxels either as leaflet tissue or background. The final iso-surface model of the leaflets is extracted from the classified voxels using the marching cube algorithm [118]. The final volumetric model of the leaflets is shown in Figure 4.5.

3.6.3. Experimental Results

The accuracy of the proposed method was evaluated using 536 singe-phase CT data sets. The data sets comprise a variety of cardiovascular diseases and due to different acquisition protocols they have heterogeneous image quality, resolution and sizes. The ground-truth for training and testing was obtained through an incremental annotation process guided by experts, which include the manual placement of anatomical landmarks and delineation of the aortic valve root surface. The volumetric models of the leaflets, calcification and blood pool are segmented using a semi-automatic process. Our data set was splitted into 408 training and 128 test data sets. The reported numbers were performed on the test data set.

First the landmark model m estimation error is computed for the most important landmarks for the TAVI procedure, the commissures and hinge points. Euclidean distances of the detected landmark points is compared to the expert annotation. The aortic root model performance, containing a dense surface mesh M, is measured as the mesh-to-mesh distance. Results shown in Table 3.9 and qualitative evaluations in Figure 4.5 corroborate that our estimation method works reliably and with high accuracy, even on low contrasted CT volumes. Within six seconds a complete personalized volumetric model of the aortic valve

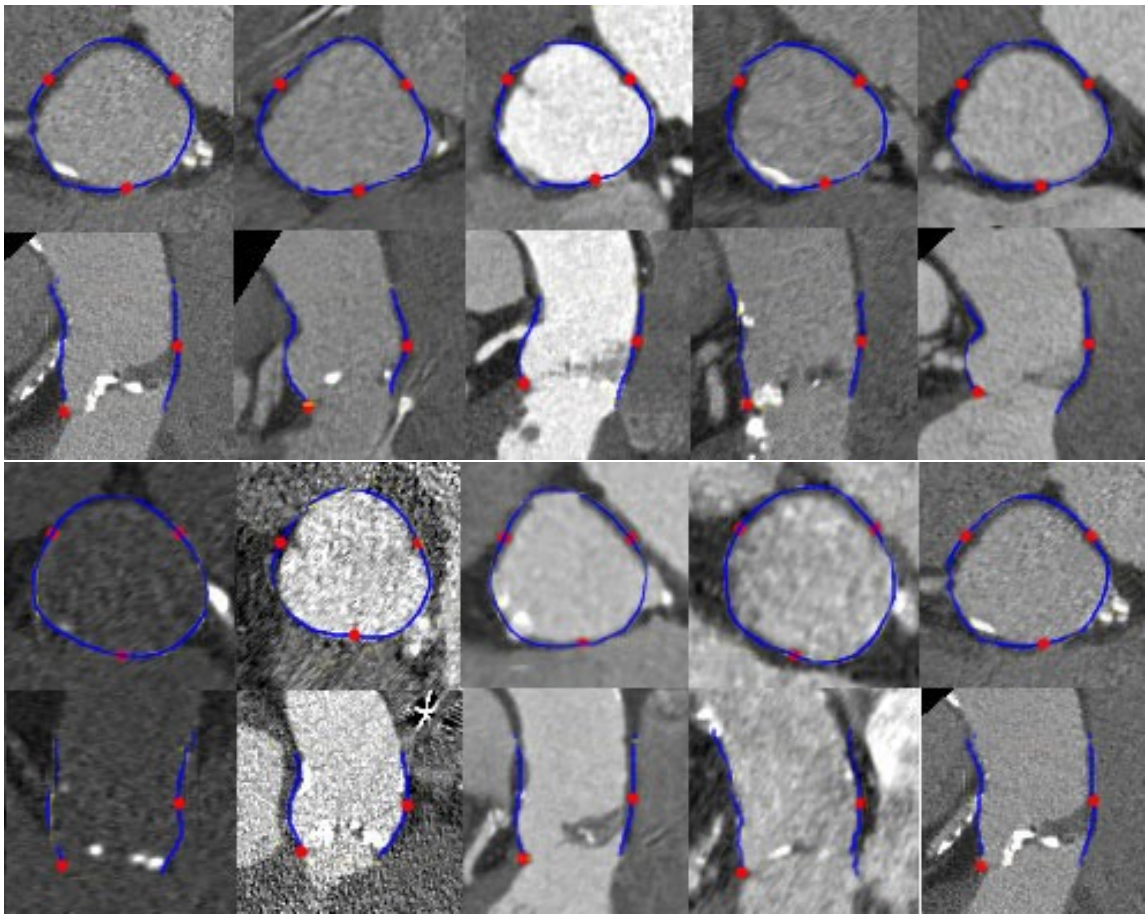


Figure 3.29.: Examples of the automatic estimation of the aortic valve surface model M for nine patient data sets. The upper row is showing short axis views of the segmentation results and the bottom long axis views.

is estimated with an average accuracy of 1.30 mm.

The viewing orientation of a C-arm system is characterized by a primary and secondary anatomic view angle. The primary angle (RAO/LAO) reports how much a C-arm has been rotated to a patient's right/left side (right anterior oblique/left anterior oblique). The secondary angle (CRAN/CAUD) tells how much a C-arm has been angulated toward a patients head (cranial) or feet (caudal) direction. In our experiments shown in table 3.9 we compared the CRAN/CAUD deviation at RAO/LAO=0 between our automated estimation and ground-truth data. Based on the hinge plane which is defined by the three aortic valve hinge points, the CRAN/CAUD angle can be computed automatically from our model. In table 3.9 the automated angulation results are compared to the ground-truth.

One important clinical measurement can be extracted from the personalized model: aortic root annulus radius. This parameter is consistently used by the physicians to determine the proper size of the artificial aortic valve implant. The Bland Altman plot shown in Figure 3.30 demonstrates a strong correlation between the measurements extracted from our personalized model M and the ground truth.

Table 3.6.: System precision for the aortic valve model estimation. The evaluation is based on landmark errors, aortic valve angulation error measured in C-arm machine coordinates and full surface mesh-to-mesh error of the aortic root.

	Mean	STD	Median
AV commissures [mm]	2.08	2.21	1.79
AV hinges [mm]	2.06	0.68	2.04
CRAN/CAUD, RAO/LAO=0 [mm]	3.98	3.47	2.90
AV root surface [mm]	1.30	0.31	1.20

The final validation is done by comparing the accuracy of our multi-class tissue classification with the ground-truth annotations. We compute the Dice similarity coefficient $DSC(A,B) = \frac{2|A \cap B|}{|A|+|B|}$ between the voxels inside the obtained segmentation and voxels inside the manual segmentation for the leaflet tissue (Figure 3.30 middle) and the calcification (Figure 3.30 right). We plot the Dice scores in respect to the tissue volume size in Figure 3.30.

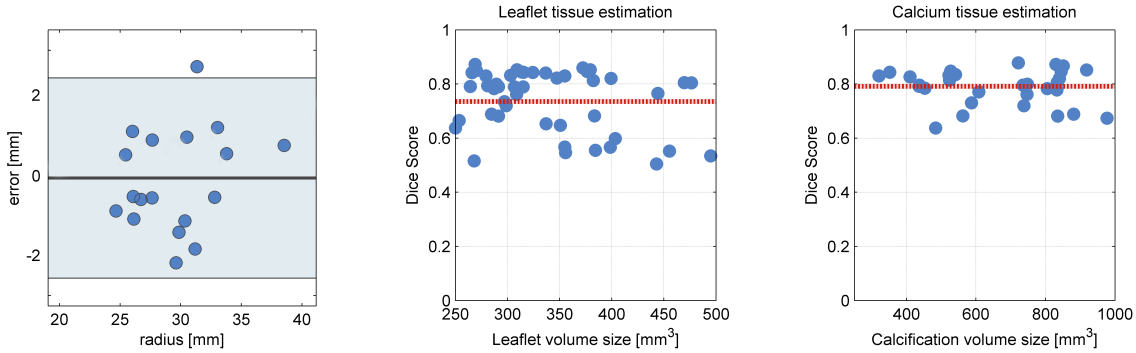


Figure 3.30.: Bland Altman plots for: Left: aortic root annulus radius. Middle: Dice score for the aortic valve leaflet segmentation ($DSC = 0.73$). Left: Dice score for the calcification segmentation inside the aortic valve ($DSC = 0.79$).

3.7. Biomechanical Modeling for TAVI planning

The most important decision during TAVI planning is selecting the proper implant type and size. Due to the wide variety in device sizes and types and non-circular annulus shapes, there is often no obvious choice for the specific patient. Most clinicians base their final decision on their previous experience. As a first step towards a more predictive planning, we propose an integrated method to estimate the aortic apparatus from CT images and compute implant deployment.

Starting from a clinical pre-operative 3D CT image (Figure 3.33), we automatically segment the aortic valve model using machine-learning algorithms and generate a patient-specific anatomical model of the aortic valve suitable for simulations (Section 3.7.1). We

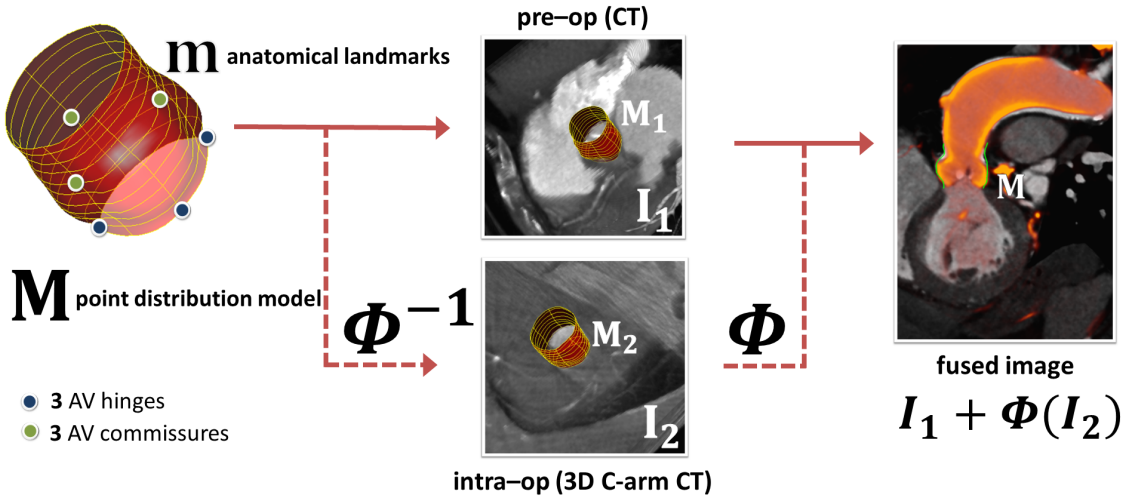


Figure 3.31.: Diagram of the problem formulation showing the surface model M, anatomical landmarks m, transformation ϕ to map the intra-op image I_2 to the pre-op data I_1.

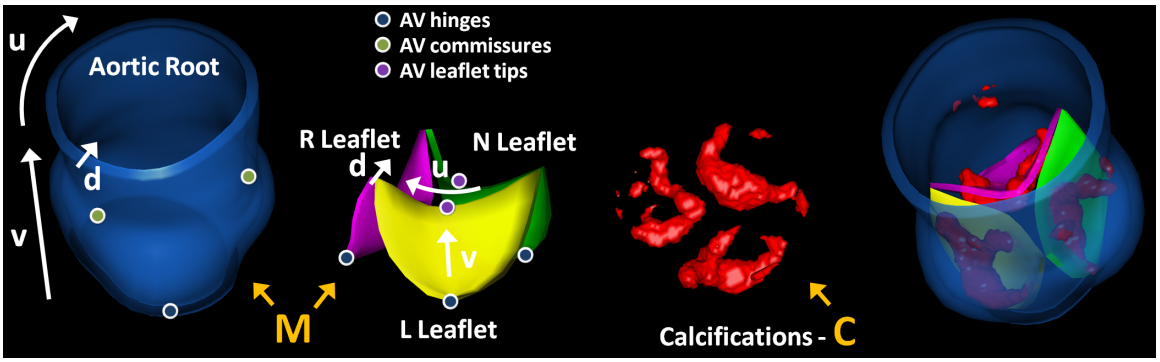


Figure 3.32.: Diagram showing the estimation framework for the volumetric aortic valve model, which consists of 9 landmarks m (3 commissures, 3 hinges and 3 leaflet tips), aortic root and aortic leaflet volumetric models M, and aortic valve calcifications C.

then apply the biomechanical model of the valve and the CoreValve implant (Section 3.7.2) to compute device deployment for TAVI planning. Finally we compare the computed geometry of the deployed implant with a ground-truth annotation extracted from the post-operative CT.

3.7.1. Parametrization and Estimation of Aortic Valve Morphology

Aortic Model Parametrization We propose a physiological and volumetric model of the aortic valve capable to capture complex morphological and pathological variations. The anatomical structures consist of nine landmarks including three commissures, three hinges,

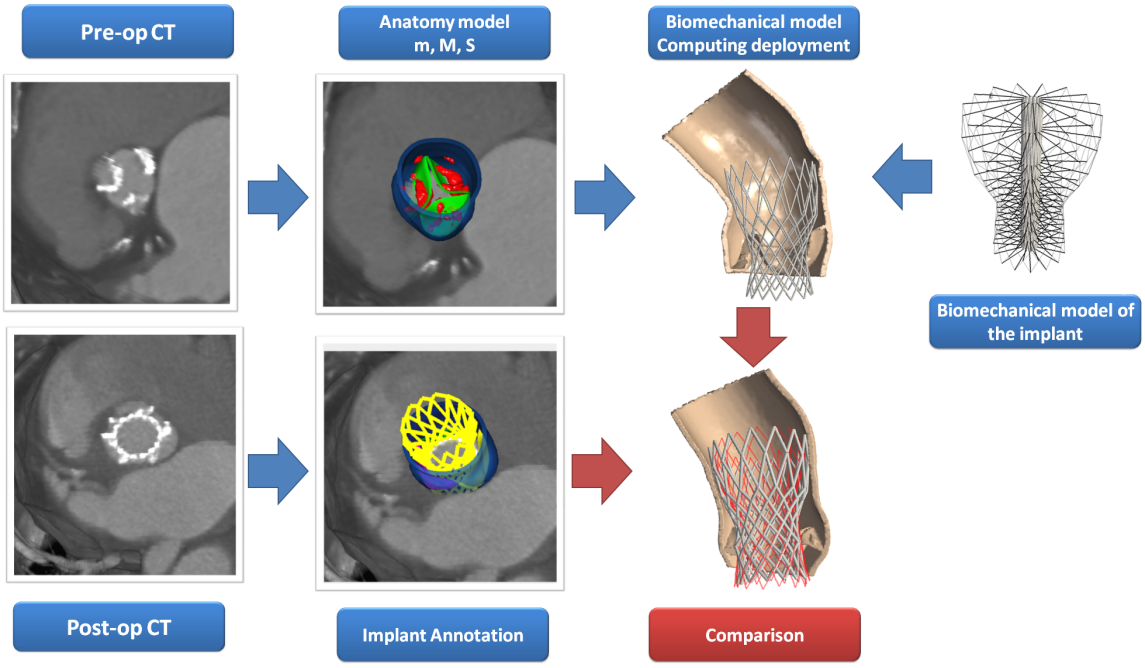


Figure 3.33.: Diagram showing our validation framework.

three leaflet tips, the aortic root, aortic leaflets and calcifications. To efficiently handle the anatomical complexity, the model representation and corresponding parametrization is constructed hierarchically using 1) a non-rigid landmark model m and 2) two volumetric models M and C for the anatomy structure and calcifications respectively (Figure 3.32). We utilize the same approach as mentioned in 3.6 to parametrize the model and estimated the model parameters.

Implant Parametrization The CoreValve implant is modeled as a tubular mesh grid aligned along the circumferential u and longitudinal v direction. We utilize a manual annotation framework to fit the ground-truth implant model to the post-operative CT data.

3.7.2. Finite Element Model of the Aortic Valve and CoreValve

From the segmentation of M and C, we build a volumetric, tetrahedral mesh whose elements are automatically tagged according to the structure it belongs to (aortic root, leaflets, calcification). The model is cut at the inflection of the aorta for computational efficiency (Fig. 3.34). The volumetric mesh is then used to compute the deformation of the aortic apparatus induced by the deploying implant. To that end, we solve the dynamics system:

$$M\ddot{\vec{U}} + C\dot{\vec{U}} + K\vec{U} = \vec{F}_c \tag{3.38}$$

where M is the lumped mass matrix, calculated from the mass density of the tissue ($\rho_t = 1070g/L$) and of the implant ($\rho_s = 6450g/L$), K is the stiffness matrix, encoding material

properties, C is a Rayleigh damping, whose coefficients for mass and stiffness matrix are both equal to 0.1, $\vec{U}$ is the displacement of the nodes of the objects in the scene and $\vec{F}_c$ is the vector gathering the external forces resulting from self-collisions and collisions with the implant stent.

Tissue Model The aortic tissue is modeled with linear isotropic elasticity for computational efficiency, although our framework can easily accommodate for more realistic hyperelastic constitutive laws that better capture the non-linear behavior of aortic tissues. Poisson ratio ν and Young's moduli E are defined per tissue type: Aortic root: $E = 2\,MPa$, $\nu = 0.48$; Aortic leaflets: $E = 1\,MPa$, $\nu = 0.48$; Calcifications: $E = 60\,GPa$, $\nu = 0.3$ [196]. Co-rotational FEM are employed to cope with large deformations.

CoreValve Model The CoreValve implant is modeled using a spring model whose stiffness is calculated directly from the specifications of the device (Young's modulus $E = 75\,GPa$). To mimic the shape-memory deployment, the model is deformed according to springs defined between the undeployed and deployed configurations (Fig. 3.34, right panel). The stiffness of these springs, $k = 1MPa$ (determined off-line), is the minimal stiffness necessary to fully undeploy the implant when free from interactions of any neighboring structures.

Boundary Conditions The aortic apparatus is tethered to the left ventricle and the aorta. To mimic the compliance of these neighboring organs, aortic annulus and aortic root are tethered in space through springs whose stiffness is equal to $10\,MPa$ (Fig. 3.34).

Contacts Implant / valve contacts and valve self-contacts are modeled using sphere contact models. For instance, for each vertex of the aortic root, a sphere of radius of 1 mm is defined. As soon as a vertex of any object (implant / valve) enters the area defined by this sphere, a spring of stiffness $100\,kPa$ is added between the two vertices to avoid contact. A contact friction of 0.1 is assumed based on [196]. The contribution of the contact forces are gathered into the contact force vector $\vec{F}_c$.

Implementation The model is implemented using the SOFA framework[1]. Spatial discretization is done using linear tetrahedra. An implicit Euler scheme is employed for time integration as it is unconditionally stable.

3.7.3. Experimental Results

3.7.4. Evaluation of Implant Deployment Prediction

We then evaluated the complete framework on eleven patients, for whom pre- and post-operative CT images where available. For these patients, the post-operative valve anatomy and CoreValve implant was obtained through an incremental annotation process guided by experts, which included manual placement of anatomical landmarks and delineation of the valve surface and implant struts model. The preoperative anatomy was detected automatically, corrected by experts if needed, and meshed with an average tetrahedral edge-length of ≈ 1.2 mm. For all patients, nominal tissue and implant parameters (Sec. 3.7.2)

[1] www.sofa-framework.org

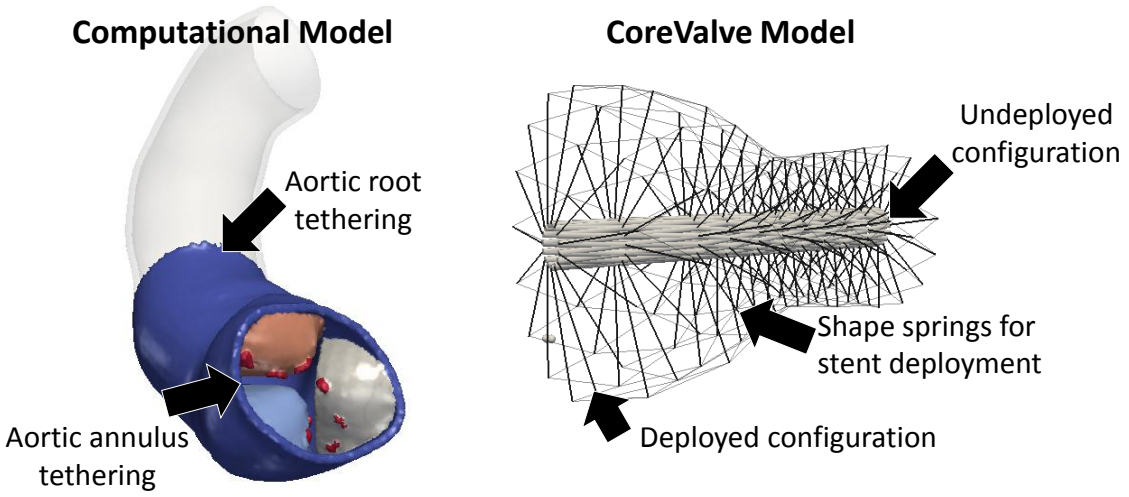


Figure 3.34.: Left: anatomical model estimated from the images. Colors encode the different anatomical parts, in red are the calcifications. Arrows indicate spatial tethering. Right: CoreValve model. Thick black lines represent the strings used to model shape-memory deployment.

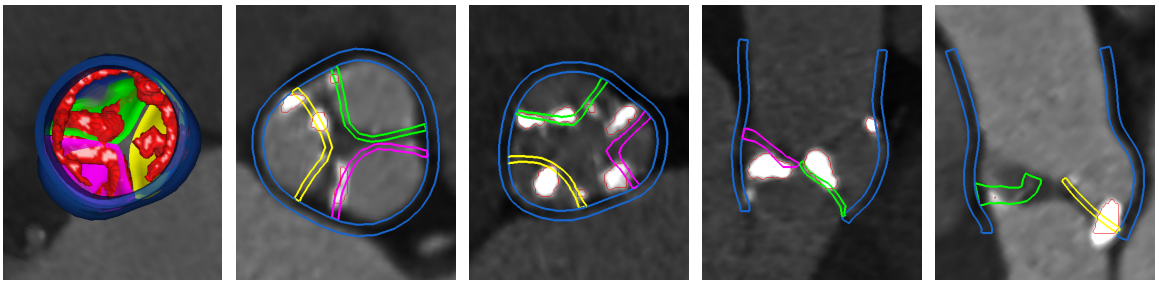


Figure 3.35.: Automatic segmentation results of the aortic valve model M and calcification C (red color) from pre-operative CT. The aortic root is shown in blue and the R-, N-, L- leaflet are shown in green, purple and red respectively. The extracted models are accurately delineating the valve anatomy in the CT image.

were employed to evaluate the robustness of the predictions with respect to tissue properties. Simulation time step was $dt = 1\,ms$. The simulation was stopped when the overall system reached an equilibrium.

Nine out of eleven cases had the aortic leaflets closed in the preoperative image. To place the virtual CoreValve device, we artificially opened the leaflets by applying a pressure of $80\,mmHg$ to their ventricular surface. Tissue and simulation parameters were kept unchanged. Then, once the valve was open, the undeployed CoreValve was placed according to the postoperative anatomy in order to reproduce as close as possible the real intervention. More precisely, the post-to-pre rigid transformation T between the aortic root model $M_{root_{post}}$ and $M_{root_{pre}}$ was estimated using Procrustes alignment. Based on that transformation, the undeployed implant model was mapped to the pre-operative data

with the same relative position as in the post-op data. The postoperative implant was also registered to the preoperative data for evaluation purposes.

Figure 3.36 reports simulation results on four patients. Despite the nominal tissue parameters and the relative simplicity of the biomechanical model, the predicted deployed implants were in close agreement with the actual postoperative outcome. More quantitatively, the point-to-mesh error between the computed deployed implant struts and the ground-truth implant annotation was of 1.73 ± 0.40 mm (Table 3.7). In addition we compared implant diameter errors for several key locations along the tubular implant structure (Table 3.8). As one can see, we could achieve precise results (1.32 mm error) for the annular ring (ring ID 6), which is the critical area for TAVI intervention. Compared to the usual implant gap of 3 mm between implant sizes, our precision has significant accuracy in clinical practice.

Table 3.7.: Point-to-mesh distance between computed and ground-truth implant deployment configuration. The average error is $1.73 \pm 0.40\,mm$.

Patient	Point-to-Mesh Error	Patient	Point-to-Mesh Error
01	$1.77 \pm 1.36\,mm$	05	$2.11 \pm 1.57\,mm$
09	$1.00 \pm 0.85\,mm$	02	$1.86 \pm 1.39\,mm$
06	$1.54 \pm 1.08\,mm$	10	$1.76 \pm 1.30\,mm$
03	$2.36 \pm 1.62\,mm$	07	$1.90 \pm 1.26\,mm$
11	$1.21 \pm 0.94\,mm$	04	$1.44 \pm 1.31\,mm$
08	$2.07 \pm 1.69\,mm$		

3.8. Interventional model estimation

With the new generation of interventional C-arm machines a 3D CT like modality can be acquired in order to guide minimally invasive procedures. In addition several protocols have been developed to acquire the vascular system using contrast agent [94]. One can directly segment the aortic valve in the contrasted image [206]. However due to the limited image quality and limited field of view the segmentation is usually less accurate compared to models extracted from pre-operative modalities. By utilizing a method which estimates the model in the intra-operative modality and pre-operative modality at the same time the interventional model accuracy can be improved.

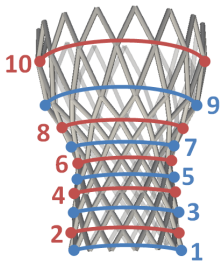
Our goal is to estimate a 3D patient-specific model M from volumetric multi-modal datasets I_1 and I_2, where I_1 is the pre-operative and I_2 the intra-operative image, and the transformation ϕ which maps the intra-operative model M_2 to the pre-operative model M_1 (see Fig 3.31).

$$(\hat{\phi}, \hat{M}) = \arg\max_{M,\phi} P(M, \phi \,|I_1, I_2\,) \tag{3.39}$$

$\phi = D\ A$ is composed of an affine- A and a non-linear warping transformation D. D is modeling the small deformation of M due to respiration and uncertainties in the acqui-

Table 3.8.: Average diameter error between computed and ground-truth implant model. Precise results were achieved in the annulus region (ring ID 6) where the error is far below the average 3mm gap between consecutive implant sizes.

Ring ID	Diameter Error	Ring ID	Diameter Error
01	$1.86 \pm 1.44\,mm$	06	$1.32 \pm 0.72\,mm$
02	$1.20 \pm 0.80\,mm$	07	$2.77 \pm 1.21\,mm$
03	$1.04 \pm 0.55\,mm$	08	$4.10 \pm 1.96\,mm$
04	$0.80 \pm 0.58\,mm$	09	$2.68 \pm 2.34\,mm$
05	$1.00 \pm 0.84\,mm$	10	$3.60 \pm 2.56\,mm$



sition phase between the pre- and intra-operative data. The model M is represented as a point distribution model. Using the transformation ϕ the pre- M_1 and intra-operative M_2 models can be computed: $M = M_1$, $M = D\ A\ M_2$ and $M_2 = A^{-1}\ D^{-1}\ M$.

3.8.1. Method

In general, finding an optimal solution to Eqn. (3.39) is difficult and has high computational cost if we want to model complex anatomical structure such as in section 3.1. Therefore we approximate the problem by expanding the formulation and exploiting mutual independence. In addition a shape constraint term is added to restrict the estimated model M in a shape space built from a database of annotations.

$$(\hat{\phi}, \hat{M}) = \arg\max_{M,\phi} \log\left(P(M\,|I_1\,) \cdot P(M\,|\phi(I_2)\,) \cdot P(M\,|I_1, \phi(I_2)\,) \cdot P(M, \phi\,|\mu, \Sigma)\,)\right) \tag{3.40}$$

We infer all the probabilities in our formulations using robust learning based algorithms. The first $P(M\,|I_1\,)$ and the second term $P(M\,|\phi(I_2)$ define the independent model estimations in the multi-modal images I_1 and I_2. As proposed in [88] a classifier is trained using the probabilistic boosting tree and Haar-features to estimate the posterior probability. The best model parameters for M are selected based on a joint probability term $P(M\,|I_1, \phi(I_2)\,)$ explained in chapter 3.8.2. The transformation ϕ is modeled as a warping transform with Gaussian radial basis functions. The last term $P(M, \phi\,|\mu, \Sigma\,)$ symbolizes a regularization of the shape M and the transformation ϕ based on the learned statistical shape model defined as a Gaussian distribution with the mean μ and the covariance matrix Σ learned from manual annotations. Both the affine A and the non-linear transformation D are updated in this stage. A bias is applied towards the pre-operative model $M = M_1$ as the model estimation is more robust in the pre-operative images. In our case I_1 represents the CT image and I_2 the TEE and 3D C-arm CT image. Our aortic valve model is parametrized as described in section 3.1.

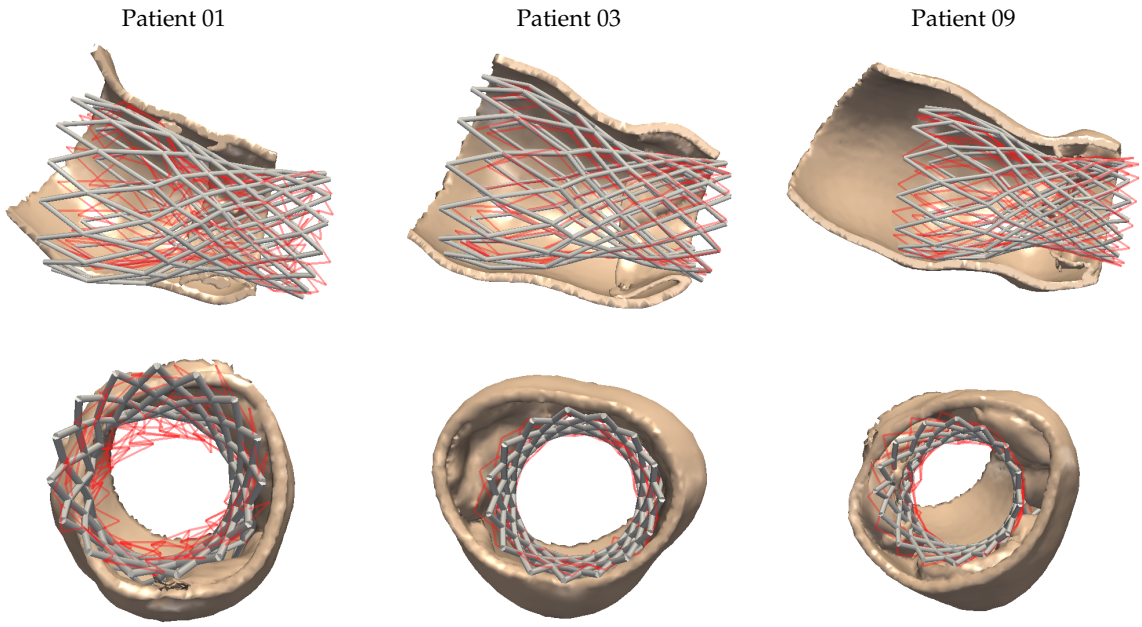


Figure 3.36.: Example of simulated implant deployment using our automatic volumetric model estimation and our simulation framework. In transparent red is the ground truth. Our model could predict CoreValve deployment based on pre-operative image data only.

3.8.2. Similarity learning

The joint term $P(M\,|I_1, \phi(I_2))$ in Eqn. 3.40 exploits the similarities between the models from the multi-modality images. Similarity functions proposed in the current literature, such as mutual information or cross correlation, could be used but as mentioned in [22] learning the similarity for a specific problem yields better performance.

In our case, we selected to use the LogitBoost [52] algorithm instead of the PBT. As mentioned in [52] the LogitBoost algorithm (see Figure 3.38) minimizes the negative binomial log-likelihood. In our context a data point x is an image pair $x = (I_1, \phi(I_2))$. Transfered to an boosted similarity score the goal of the classifier similarity function $\boldsymbol{S}(I_1, \phi(I_2))$ is the probability of the class label $y(I_1, \phi(I_2))$ having the posterior probability of 1, that is $\boldsymbol{S}(I_1, \phi(I_2)) = p(I_1, \phi(I_2))$.

We employ the Logitboost framework in order to train a cascade of strong classifiers. Each strong classifier $\boldsymbol{F}_{strong}$ consists of k weak classifiers $\boldsymbol{F}_{weak}$ which learn the similarity between pairs of image patches $I_{S1} \in I_1$ and $I_{S2} \in I_2$, $\boldsymbol{F}_{weak}(I_{S1}, I_{S2})$. The weak learners are constructed based on Haar-like features [145, 147] extracted locally from rectangular patches I_{S1} and I_{S2} from image slices sampled perpendicular to the tubular aortic root surfaces M_1 and M_2. We parametrize the rectangle feature g by (r, c, dr, dc, t) where (r, c) is the starting point of the rectangle, (dr, dc) is the height and width, and t is the feature type. There are six feature types as shown in Figure 3.39. Given a rectangle feature g and

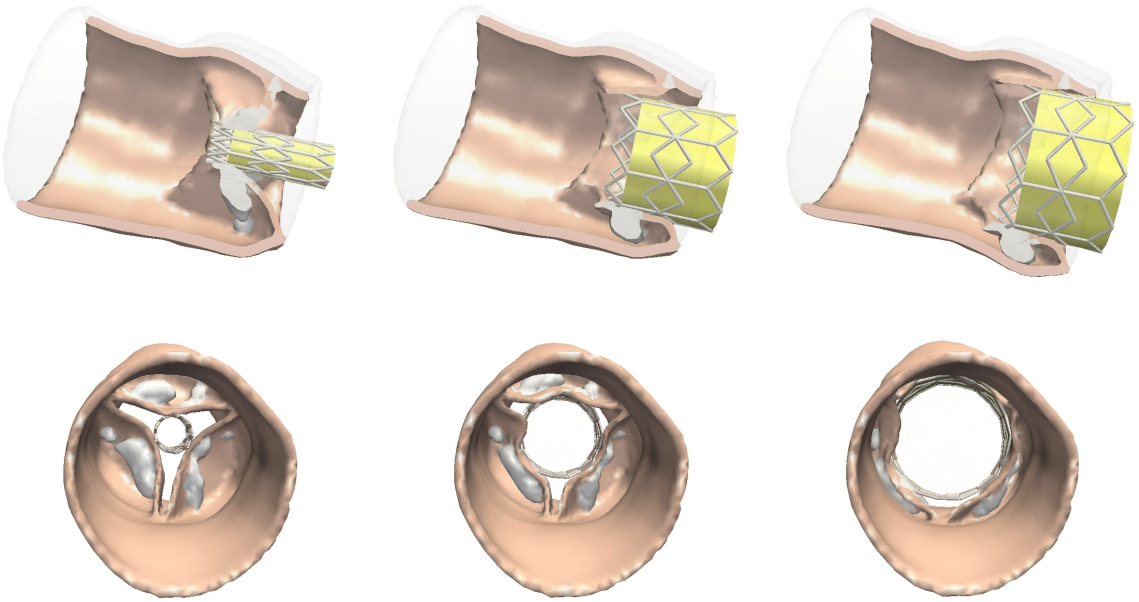


Figure 3.37.: Example of simulated stent deployment using our automatic volumetric model estimation and a standard machanical simulation framework. As we only estimate the volumetric model of the aortic valve leaflets a standard non-patient specific thickness of 1.4mm was assigned to the aortic valve root.

the image pair (I_1, I_2), we compute two feature responses $g(I_1)$ and $g(I_2)$ from the two integral images associated with I_1 and I_2 respectively. The two local rectangles could have different parameters; however we refrain from doing this because empirically this shows no clear advantage but significantly increases training complexity.

The weak learner is modeled as a 2D piecewise constant function defined on a 2D feature space by the feature responses of $h(I_{S1})$ and $h(I_{S1})$. The 2D feature space is separated in equal rectangular non-overlapping regions. Therefore we quantize the feature responses from both modalities in 64×64 bins whereby the values are scaled between the minimum and maximum feature responses $h(I_{S1})$ and $h(I_{S1})$.

$$\boldsymbol{F}_{weak}(I_{S1}, I_{S2}) = \sum_{b=1}^{B} \sum_{c=1}^{C} \beta_{b,c} \, R_{b,c} \left[h(I_{S1}) \times h(I_{S2}) \right] \tag{3.41}$$

where B and C are the bin numbers for the feature responses in each modality and $\beta_{b,c}$ symbolizes the constant associated with the region $R_{b,c}$ representing a bin in the 2D feature space. As in [208] the optimal weights $\beta_{b,c}$ would be determined by fitting a least-squares regression function. During detection a probability for each weak classifier is evaluated by extracting Haar-features from pairs of image patches. The features are assigned to a bin $R_{b,c}$ based on the feature response and multiplied with the corresponding weight $\beta_{b,c}$. We empirically determine the interval boundary points by uniformly dividing the feature

Input

- set of N labeled samples $\{(x_1, y_1), ..., (x_N, y_N)\}$
- distribution W over the examples
- weak learning algorithm **WeakLearn**
- number of iterations T

Initialization

- set the weight vector $w_i^l = W(i) = 1/N$ for $i = 1, ..., N$ and the probability estimates $p(x_i) = 1/2$.

Main Loop, for t=1,...,T

- compute working responses and weights:

$$z_i = \frac{y_i - p(x_i)}{p(x_i)(1 - p(x_i))};$$
$$w_i = p(x_i)(1 - p(x_i)).$$

- fit the function $f_m(x)$ by a weighted least-squares (LS) regression of z_i to x_i with weights w_i.

$$f_m(x) = argmin_{f \in F}(\epsilon(f) = \sum_{i=1}^{N} w_i(z_i - f(x_i))^2).$$

- Update $F(x) \leftarrow F(x) + \frac{1}{2} f_m(x)$ and $p(x)$ via

$$p(x) = \frac{exp(F(x))}{exp(F(x)) + exp(-F(x))}.$$

Output: Classifier $sign[F(x)]$.

Figure 3.38.: The two-class LogitBoost algorithm as introduced by Friedman in [52].

responses.

Given a weak learner $\boldsymbol{F}_{weak}$ that is associated with a feature h, the optimal weight $\beta_{b,c}$ that minimizes the weighted LS cost f in 3.38 is the weighted response z of all data points falling into the region

$$\beta_{b,c} = \frac{\sum_{i=1}^{N} w_i \, z_i \, R_{b,c}\,[h(I_{S1}) \times h(I_{S2})]}{\sum_{i=1}^{N} w_i \, R_{b,c}\,[h(I_{S1}) \times h(I_{S2})]} \tag{3.42}$$

where I_{S1}, I_{S2} is the i-th training image pair. Figure 3.39 illustrates the fitting process. Figure 3.39 left) and center) visualizes the field of $w_i \times z_i = y_i - p(x_i) = 1 - p(xi)$ for

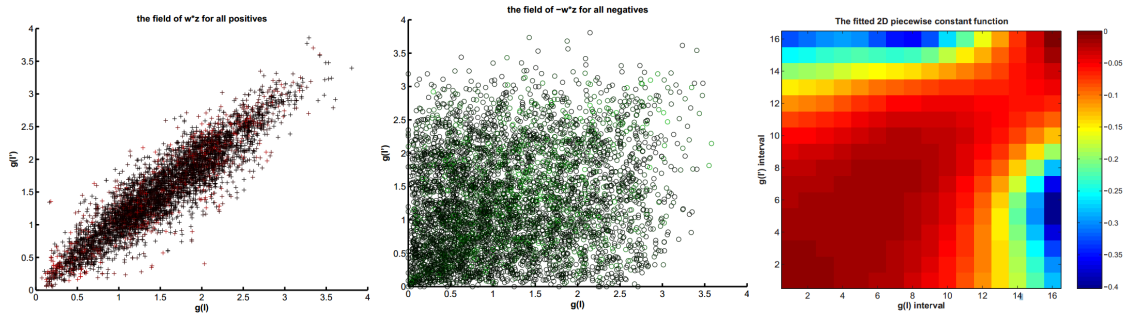


Figure 3.39.: Left and middle: illustrating the process of fitting a 2D piecewise constant function. Left: illustrates the fitted 2D piecewise constant function.

all positives, where the color intensity corresponds to the value of $w \times z$: the greener the plus sign is, the less likely the data point x is positive. The diagonal structure in Figure 3.39 shows that the two feature responses of the positives are roughly same. Figure 3.39 visualizes the field of $w_i \times z_i = p(x_i)$ for all negatives: the greener the circle sign is, the less likely the data point x is negative. As shown in Figure 3.39 middle), the negatives are characterized by a widely-dispersed nature. It shows the fitted 2D PWC function: the constant coefficients $\beta_{b,c}$ along the diagonal lines are high, while offdiagonal ones are low. For the step 2 in Figure 3.38, the weak function f with the smallest weighted LS cost $\epsilon(f)$ is selected.

The use of nonparametric 2D PWC functions as weak learners is beneficial. Take the 1D case for example; 1D simple regression stumps that binarize the feature response are often used as weak learners in the literature [176] . It is possible to verify that any 1D PWC function can be constructed by combining multiple 1D simple regression stumps. The similar holds for the 2D case. Such a combination strengthens the modeling power of weak learners and consequently accelerates the training process. Empirical evidence shows that the learning time is almost inversely proportional to the number of thresholds used in the weak learner. One may argue that it brings the risk of overfitting. But boosting combats the problem overfitting in terms of classification even when the weak learner overfits. Further, in practice we smooth the fields of $w \times z$ and w before taking the division to ameliorate the overfitting of the weak learner itself.

A cascade of l strong classifiers $\boldsymbol{F}_{strong}$ is trained in order to determine the posterior probability $P(M\,|I_1, \phi(I_2)) = \boldsymbol{S}(I_{S1}, I_{S2})$ of the similarity function.

3.8.3. Model-Based Fusion Approach

The first stage in our hierarchical model estimation algorithm consists of pre-aligning the multi-modal images using the anatomical landmarks. The affine transformation A is estimated by obtaining a least-squares solution based on the independently detected landmarks m_1 from the image I_1 and m_2 from the image I_2. The landmark detectors are trained using the PBT classifier and Haar-like features. The surface M is initialized by learning a

correlation model between measurements extracted from the landmarks m_1 and the point distribution model M, as described in [67]. The nonlinear warping transformation D is set to identity. Based on A the model M can be projected to the image I_2.

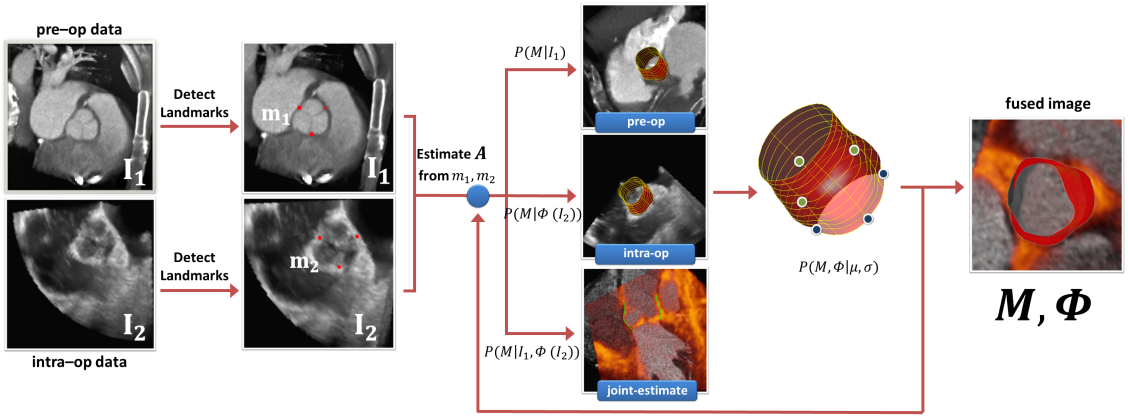


Figure 3.40.: Diagram showing the model based fusion approach for the estimation of the model M and the transformation ϕ.

In the optimization phase we apply an iterative approach. We sample candidates N_1 and N_2 along the surfaces normals of M_1 and M_2, and evaluate the probability $P(M\,|I_1\,)$ for each candidate $n_1 \in N_1$ and $P(M\,|\phi(I_2)\,)$ for each point $n_2 \in N_2$ using the learned detectors. The joint probability $P(M\,|I_1, \phi(I_2)\,)$ is determined by training a boosting classifier, as in Section 3.8.2, to evaluate pairs of candidates. A cross product of the candidates $N_1 \times N_2$ is constructed and the highest probable candidate pair (n_i, n_j) is selected by multiplying the single modality probabilities with the joint term.

$$(n_i, n_j) = \arg\max_{n_i, n_j} \left(P(n_i\,|I_1\,) \cdot P(n_j\,|\phi(I_2)\,) \cdot P(n_i, n_j\,|I_1, \phi(I_2)\,)\right) \tag{3.43}$$

The estimated candidate pairs are used to update the models M_1 and M_2. The second step of the iteration involves calculating the posterior probability $P(M, \phi\,|\mu, \Sigma)$ of M and ϕ based on the learned statistical shape models. This could be perceived as a regularization to the shape M. Thereby M_1 is projected to the PCA shape space using the largest 40 eigenvectors. ϕ is updated by computing the rigid transformation R based on the posterior probability of the pairs (n_i, n_j). D is updated by obtaining a least-squares solution to the warping transformation $\hat{D} = \arg\min \|T\ M_2 - D^{-1} M_1\|^2$ using radial basis functions. Thereby the number of control points is much smaller than the number of shape points M. The algorithm converges in a small number of iterations. Figure 3.40 demonstrates the complete estimation approach.

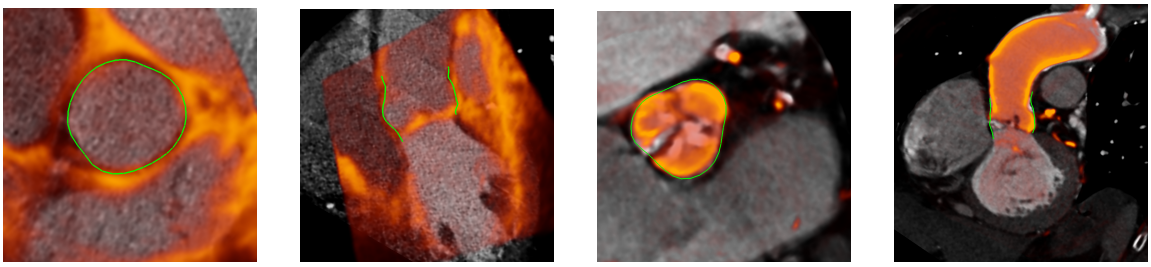


Figure 3.41.: Example of the joint aortic valve model estimation from pre- and intra-op volumetric data. The left 2 images show fused CT-TEE data sets and the right 2 images show fused CT-3D C-arm CT data. The mapping of the intra-op image I_2 to the pre-op image I_1 is done by the estimated non-linear transform ϕ.

3.8.4. Experimental Results

The most relevant intra-operative modalities with 3D capabilities (3D C-arm CT and TEE) in the OR environment were incorporated for evaluation. In total 56 volumes, 13 pairs of CT-TEE data sets and 15 pairs of CT-3D C-arm CT data pairs were selected to demonstrate the effectiveness of our method. This dataset was solely used for evaluation and not included in training. The ground-truth annotations were obtained from clinical experts by manually placing the anatomical landmarks in the pre- (m_1) and intra-op (m_2) images and finally delineating the aortic valve surfaces M_1 and M_2.

As our algorithm depends on the automatic detection of the anatomical landmarks m_1 and m_2 during the initialization step in order to estimate the affine transform A we evaluate their detection performance on the test dataset. For training 160 separate landmarks annotations in CT, 320 in TEE and 192 in 3D C-arm CT were used to train the landmark detectors. The error is computed as the Euclidean distance between the automatic estimation and the expert annotation. For the hinges we obtain an error of $2.40 \pm 0.81mm$ in CT, $2.56 \pm 0.71mm$ in TEE and $2.30 \pm 1.56mm$ in 3D C-arm CT and for the commissures $2.74 \pm 1.01mm$ in CT, $3.31 \pm 1.55mm$ in TEE and $2.98 \pm 1.44mm$ in 3D C-arm CT. The results are shown in Table 3.9.

Table 3.9.: System precision for the estimation of aortic valve landmarks (commissures and hinges) in CT, TEE and 3D C-arm CT.

	aortic valve hinges			aortic valve commissures		
	Mean	STD	Median	Mean	STD	Median
CT [mm]	2.40	0.81	2.19	2.74	1.01	2.48
TEE [mm]	2.56	0.71	2.26	3.31	1.55	2.81
3D C-arm CT [mm]	3.59	1.17	3.47	3.79	1.02	3.82

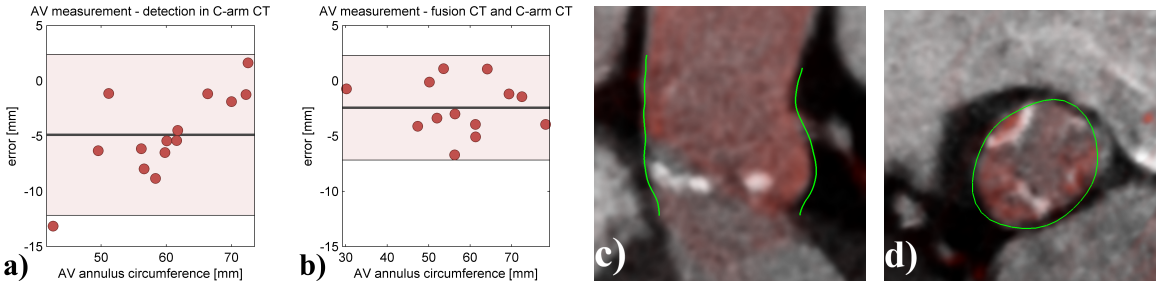
The mesh-to-mesh error was computed between the ground-truth annotations and the detected models in order to obtain quantitative results for the automatic surface estima-

tion. Results shown in Table 3.10 confirm that our model-based fusion estimation approach yields the best results among the studied approaches.

Table 3.10.: System precision for aortic valve surface model estimation in CT, TEE and 3D C-arm CT. Comparison between our novel model-based fusion approach and single modality estimations.

	single modality estimation			fusion approach		
	Mean	STD	Median	Mean	STD	Median
CT-TEE [mm]	1.22	0.23	1.13	1.09	0.22	1.10
CT-3D C-arm CT [mm]	1.96	0.54	1.99	1.73	0.49	1.79

For TAVI the selection of the appropriate stent size and its positioning have clinical significance. However in 3D C-arm CT the aortic valve annulus is not visible as the contrast is injected at the cusp area. Fusing the 3D C-arm image with pre-op CT data would allow the physician to properly examine the annulus area and enable accurate positioning of the stent during the procedure. We evaluate the error for the aortic valve annulus ring circumference, extracted from the estimated aortic valve model M, by comparing the result of the independent detection in 3D C-arm CT image and our model-based fusion approach. Quantitative and qualitative results are shown in Figure 3.42.



(a) (b) (c) (d)

Figure 3.42.: Bland-Altman plots for the aortic valve annulus circumference measurement extracted from the model M with (a) independent detection in 3D C-arm CT and (b) fusion of pre-op CT and 3D C-arm CT. (c) and (d) are showing short and long axis views of the model M and the fused pre- and intra-op images $I_1 + \phi(I_2)$.

3.9. Discussion

In this chapter we proposed several methods for quantification and pre-operative modeling for minimally invasive valve procedures. First we presented a novel personalized model for quantitative and qualitative evaluation of the complete heart valve apparatus from multi-phase CT. It is capable to delineate the full anatomy and dynamics needed to depict a large variation of valve pathologies, especially diseases affecting several valves.

Its hierarchical approach using state of the art machine learning algorithms in combination with a constrained Multi-linear shape space enables patient specific model estimation within three minutes and an accuracy of 1.24 mm. Our experiments showed that the automatic estimation of the valve models are slightly above (1.13 mm for the aortic valve and 0.78 mm for the mitral valve) the inter-user variability. Considering that our detection was done on 1mm resolution this means that the deviation is around one additional voxel apart from the variability of the expert annotations. Data sets with low contrast agent, noise in the data and low imaging quality were the main reasons for the performance gap between the automatic estimation and the inter-user variability. As in most of our data sets the contrast agent was more concentrated in the left side of the heart and thus the estimation accuracy for the pulmonary and tricuspid valve was inferior to the aortic and mitral valve.

Additionally, we proposed an extension of this method for the aortic valve in the context of pre-operative planning for TAVI. In addition to the parametrization and estimation of the aortic valve model advanced evaluations necessary for comprehensive planning such as calcification distribution assessment and tissue characteristics can be extracted from our volumetric models. Thus, our method has the potential to enable novel clinical applications to address the two major drawbacks of current TAVI procedures (paravalvular leakages and stroke) which are correlated with calcium quantity and distribution.

Afterward we proposed an integrated framework for personalized computation of TAVI deployment. Our approach enables fully automated model extraction from pre-operative CT data and patient-specific implant deployment simulations. We have demonstrated the validity our framework to predict post-operative implant geometry on eleven patients undergoing TAVI with pre- and post-operative data.

Finally we propose a novel approach to estimate comprehensive patient specific models of the aortic valve by model-sensitive fusing of multimodal pre- and intra-operative data. Fast and robust machine learning techniques are employed during the estimation exploiting redundant and complementary information from the multimodal images. Thereby high-quality patient-specific models are integrated into the imaging environment of operating rooms to guide cardiac interventions. Comprehensive quantitative and qualitative experiments on the aortic valve modeling demonstrate the effectiveness of our approach with an accuracy of 1.09mm in CT-TEE and 1.73mm in CT-3D C-arm CT.

Future work will continue to focus on the modeling side. One important extension will be the inclusion of the subvalvular apparatus of the mitral and tricuspid valves. Although critical in the clinical context, these structures are difficult to distinguish. Thus, patient-specific parameters must be inferred from statistical models or by fusing additional imaging information. The extension to a volumetric representation for the remaining valves (MV, PV, TV), which models the tissue thickness is also of high clinical importance. Such parameters could be estimated within the same discriminative learning framework.

4. Intra-operative guidance

In the context of minimally invasive procedures, where the defected anatomy is not directly visible or accessible, guidance using different imaging modalities plays a key role to ensure an optimal procedural outcome. Most interventional procedures are performed within Hybrid operating rooms (ORs) where advanced imaging techniques, such as Fluoroscopy and TEE Ultrasound are available. Some approaches work solely on the interventional imaging data to provide intra-operative guidance [136, 79, 57, 56] however this task is extremely challenging because of the limited data quality of the intrerventional images. Fluoroscopy provides only a projection image and thus there is no depth information, soft tissue can only be seen with additional contrast injection which complicates the acquisition protocol. TEE ultrasound enables the visualization of soft tissue but its drawbacks are a limited field of view (FOV) of the acquired image and the low signal-noise ratio.

Most methods used for interventional guidance however rely on pre-operative data which is available due to the crucial planning step for interventional procedures. This data is of high quality (CT, MRI) with a large FOV. As the acquisition procedure is standardized for CT (including contrasted CT) and MRI, a high-resolution image containing both soft tissue and bones can be acquired. In addition anatomical models can be extracted from this image (see chapter 3) and used during guidance. Thus using the pre-operative images the task of interventional guidance simplifies to aligning pre-opreative images (or derived anatomy) to the lower quality interventional images. To solve this problem registration methods can be utilized.

Within the last decade new interventional C-arm devices, such as the Zeego, allow to acquire a CT-like 3D modality in addition to the standard 2D Fluoroscopy. This new modality, 3D C-arm CT, has a lower resolution compared to CT and a significantly lower signal-to-noise ratio. However it simplifies the problem of alignment from a 3D-2D to a 3D-3D problem. If alignment between the pre-operative modality and the intra-operative 3D C-arm CT is achieved both anatomical models and the pre-operative volume information can be overlayed on the 2D Flouroscopic image as both are acquired within the same coordinate system.

In the next sections the basic concept of image registration will be introduced. Furthermore, the components of a state-of-the-art registration framework will be explained followed by two novel methods to register images based on geometric models rather than image intensities.

4.1. Image-based Registration for Intra-Operative Guidance

The term "'registration"' means the establishment of a common coordinate system for different medical images. In our case the alignment of high-quality pre-operative data with limited intra-operative data can be used to guide minimally invasive procedures. Detailed introductions in the field of medical image registration can be found in [74] and in [11]. Another general review can be found in [125, 92, 151, 126, 137], where a categorization of the existing registration methods is provided. A more recent review of specific 2D-3D registration methods was presented in [128].

Depending on which information is used to align the images current registration methods can be separated into Image-based and Feature-based methods.

Image-based methods use the image intensities directly to compute the registration. Sometimes the images are automatically pre-processed before the registration process starts. In most cases these methods perform the registration automatically except for the pre-alignment which can be manual. In some other cases manual masking of the area of interest is important as a prerequisite for registration to avoid the sensitivity to local optima.

Feature based methods usually do not work on the image intensity directly. They rather extract indirect information before the registration process starts. This information could be points or surfaces. By aligning these points or surfaces the two modalities can be brought into correspondence [5]. The points and their correspondence can be determined using markers placed on the patients skin or implanted into the patients. The points can also be defined by manually marking corresponding anatomical landmarks in all images. The final alignment, in the case of a rigid alignment, can be determined in a closed-solution form [182]. In the case that surfaces are extracted (manually or automatically) methods such as ICP (iterative closest point) [203], the CPD (coherent point drift) [139, 140] or other more robust variants can be utilized [65, 142]. Many methods extract the points [183] and surfaces automatically, simplifying the process of alignment for the end-user. Recently methods utilizing machine learning techniques have been proposed in order to eliminate the manual interaction during registration [142, 102]. In the next sections we will introduce the general registration concept and its components which are valid for most volumetric based registration methods. We will then proceed to introduce our new method to fuse pre-operative and intra-operative volumetric images [142].

4.1.1. General Formulation

We consider that images are defined as scalar vector functions:

$$I : \Omega \rightarrow \mathbb{R} \tag{4.1}$$

Ω symbolizes the domain on which the image I is defined. Based on the location $x \in \Omega$ the image intensity at this point is defined as $i = I(x)$. The location x is usually expressed in a Cartesian coordinate system (in our case three dimensional) and is either measured in voxel coordinates (or physical units like mm where the spacing of the image is incor-

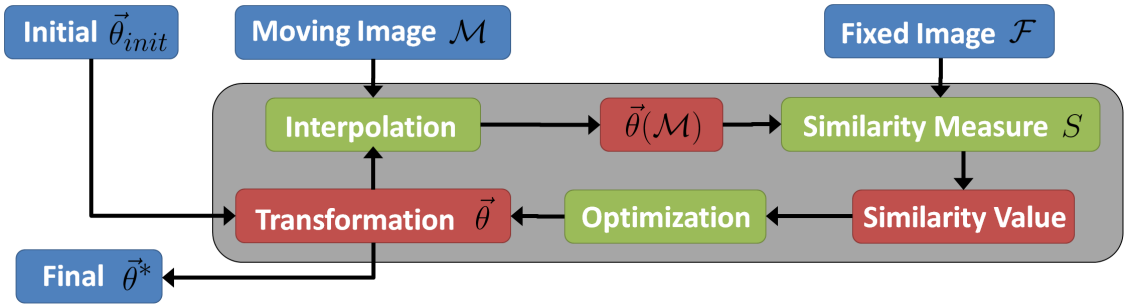


Figure 4.1.: Overview of standard iterative registration workflow.

porated). As the image represents a discrete Cartesian grid, the image intensity can be sampled only in this discrete locations. As registration algorithms need to evaluate the image at non-discrete locations interpolation is needed (see Figure 4.1). The assumption that interpolation is implicitly performed when required, allows us to use the definition of images on a continuous domain Ω.

In the case that two images $\mathcal{F}$ and $\mathcal{M}$ are given the goal of the registration algorithm is to find the optimal transformation model $\vec{\theta}$ that maps points $x_{\mathcal{M}} \in \Omega_{\mathcal{M}}$ from the image $\mathcal{M}$ into the points $x_{\mathcal{F}} = \vec{\theta}(x) \in \Omega_{\mathcal{F}}$, such that the underlying anatomy is correctly aligned. Most of the literature refers to $\mathcal{F}$ as the fixed image as its original domain is used and $\mathcal{M}$ as the moving image because its location is evaluated depending on the transformation model $\vec{\theta}$.

A registration algorithms tries to estimate the optimal parameters of the transformation $\vec{\theta}$ in order to minimize a similarity measure S.

$$\vec{\theta}^* = \arg\max_{\vec{\theta}} S(\mathcal{F}, \vec{\theta}(\mathcal{M})) \tag{4.2}$$

where $\vec{\theta}(\mathcal{M})$ symbolizes the transformation of the whole image $\mathcal{M}$. The similarity measure S assesses the quality of alignment with regard to the underlying anatomy from the image $\mathcal{F}$ and the transformed image $\mathcal{M}$. Due to the fact that the similarity measure is defined on the irregular distribution of the underlying intensity values of the images no elegant closed form solution can be derived to solve the optimization problem. Therefore, the registration is usually an iterative process, as outlined in Figure 4.1.

4.1.2. Transformation Models $\vec{\theta}$

A variety of transformation models can be used in order to transform the moving image $\mathcal{M}$ and bring it in alignment with the fixed image $\mathcal{F}$. Figure 4.2 illustrates common transformation types.

As our registration method will be utilized in the three dimensional space we start with the notion of homogenous coordinates. It is a 3D-vector extended with a fourth coordinate that is initially set to 1. In particular, it allows for the compact representation of all global

3D transformations as a 4×4 matrix. Its full strength is deployed when it comes to projective transformations, see [170] or [76] for details. Homogenous coordinates are also the standard notation in the area of computer graphics [48].

Rigid

A rigid transformation $\vec{\theta}_{Rigid}$ applies a translation and a rotation to each location on the moving image $\mathcal{M}$. This type of transformation preserves all geometric properties such as length, volume, parallelity etc. It has six degrees of freedom (DOF), three for translation and three for rotation.

$$\vec{\theta}(x) = R\,x + t \tag{4.3}$$

where $t = (t_x, t_y, t_z)^T \in \mathbb{R}^3$ is a vector describing the translation and R is a 3×3 orthonormal matrix, where its transposed equals the inverse $R^{-1} = R^T$. If homogenous coordinates are used, a 4×4 matrix can be composed containing both the translation T and rotational R parameters.

$$\vec{\theta}_{rigid} = \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{4.4}$$

The three column vectors

$$\begin{pmatrix} r_{11} \\ r_{21} \\ r_{31} \end{pmatrix}, \begin{pmatrix} r_{12} \\ r_{22} \\ r_{32} \end{pmatrix}, \begin{pmatrix} r_{13} \\ r_{23} \\ r_{33} \end{pmatrix} \tag{4.5}$$

are basis vectors of the rotated coordinate system, all have unit length and are orthogonal to each other.

Parametrization of R using Euler Angles

An intuitive and common representation of 3D rotations are Euler angles, where the rotation R is decomposed into three parts: rotation around the x-axis (ϕ_x), y-axis (ϕ_y) and z-axis (ϕ_z).

Given a set of rotation angles ϕ_x, ϕ_y, ϕ_z the final rotation matrix R can be composed as follows:

$$R = R_x(\phi_x) \times R_y(\phi_y) \times R_z(\phi_z) \tag{4.6}$$

where $R_x(\phi_x)$, $R_y(\phi_y)$ and $R_z(\phi_z)$ are defined as follows

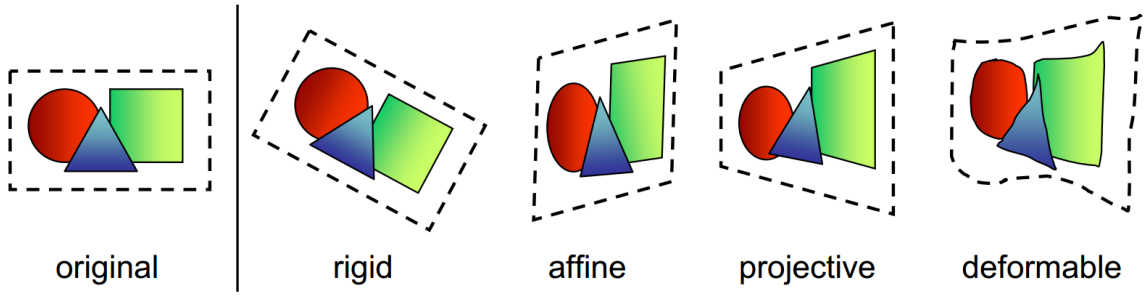


Figure 4.2.: Overview of common transformation models applied to a 2D object.

$$R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi_x) & \sin(\phi_x) \\ 0 & -\sin(\phi_x) & \cos(\phi_x) \end{pmatrix} \quad R_y = \begin{pmatrix} \cos(\phi_y) & 0 & \sin(\phi_y) \\ 0 & 1 & 0 \\ -\sin(\phi_y) & 0 & \cos(\phi_y) \end{pmatrix} \quad R_z = \begin{pmatrix} \cos(\phi_z) & -\sin(\phi_z) & 0 \\ \sin(\phi_z) & \cos(\phi_z) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

If a rotation matrix R is given the Euler angles can be computed as follows

$$\phi_x = \tan^{-1}\left(\frac{r_{23}}{r_{33}}\right)$$

$$\phi_y = \tan^{-1}\left(r_{31}\right)$$

$$\phi_z = \tan^{-1}\left(\frac{r_{12}}{r_{11}}\right)$$

where r_{ij} describes the i-th row and j-th column in the rotation matrix R. Another more robust way to compute the angles is the following:

$$\phi_z = \tan^{-1} R\left(\frac{r_{12}}{r_{11}}\right)$$

$$\phi_x = \tan^{-1} R\left(\frac{r_{31}\sin(\phi_z) - r_{32}\cos(\phi_z)}{r_{22}\cos(\phi_z) - r_{21}\sin(\phi_z)}\right)$$

$$\phi_y = \tan^{-1} R\left(\frac{-r_{13}}{r_{11}\cos(\phi_z) - r_{12}\sin(\phi_z)}\right)$$

Hereby only one type of inverse trigonometric operations ($\tan^{-1}$) is used and the first computed Euler angle ϕ_z is used to derive the other two.

Parametrization of R using Quaternion

Quaternions, also known as versors, are a common technique to represent orientations in 3D space. It has some advantages in comparison to Euler angles and rotational matrices.By using quaternions the gimbal lock, which defines a degenerated state of the rotation where

two of the three axes are driven into parallel configuration, can be avoided. In contrast to rotation matrices they provide better numerical stability. Thus a rigid transformation can be represented as a vector $[t_x, t_y, t_z, q_x, q_y, q_z, q_w]^T$, combining the three translational parameters and the four elements of the quaternion, describing a rotation in 3D space. A quaternion q is defined as follows

$$q = iq_x + jq_y + kq_z + q_2$$

where the values of i, j and k have the following properties

$$i^2 = -1,\ j^2 = -1,\ k^2 = -1$$

$$ij = k,\ ji = -k,\ jk = i,\ kj = -i,\ ki = j,\ ik = -j$$

Arithmetic operations associated with quaternions are similar to arithmetic operations associated with complex numbers [84, 2]. Quaternions have to be normalized:

$$q_x^2 + q_y^2 + q_z^2 + q_w^2 = 1$$

$$q_x^{'} = \frac{q_x}{|Q|}\ q_y^{'} = \frac{q_y}{|Q|}\ q_z^{'} = \frac{q_z}{|Q|}\ q_w^{'} = \frac{q_w}{|Q|}$$

$$|Q| = \sqrt{q_x^2 + q_y^2 + q_z^2 + q_w^2}$$

If the quaternion is normalized the quaternion can been seen as a rotation around a specific axis. Hereby the vector $(q_x^{'}, q_y^{'}, q_z^{'})$ represents the axis of rotation and ϕ represents the angle, calculated as $\phi = 2\cos^{-1}(q_w^{'})$.

4.1.3. Affine Transformation

Affine transformation contain in addition to a rigid transformation shearing and non-uniform scaling.

$$A(x) = HSRx + t \tag{4.7}$$

while H and S stand for

$$H = \begin{pmatrix} 1 & h_{xy} & h_{xz} \\ 0 & 1 & h_{yz} \\ 0 & 0 & 1 \end{pmatrix} \quad S = \begin{pmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & s_z \end{pmatrix} \tag{4.8}$$

H represents a shearing matrix with three components: h_{xy}, h_{yz} and h_{xz}. h_{xy} defines the shearing factor for the $x - y$ dimension, h_{yz} in the $y - z$ dimension and h_{xz} in the $x - z$ dimension. Thus the final affine transformation in 3D has 12 parameters (translation, rotation, scaling and shearing). The final representation in the homogeneous form is defined

as

$$Q = HSR; \qquad A_{affine} = \begin{pmatrix} Q & t \\ 0 & 1 \end{pmatrix} \tag{4.9}$$

4.1.4. Projective Transformation

In 3D, a projective transformation would map points from the 3D space to a 2D location. The main application in the medical imaging domain of this transformation is 2D X-ray Fluoroscopy, acquired during an interventional procedure. In most cases a mapping from a pre-operative 3D modality (CT, MRI etc.) is defined through a 4×4 projection matrix. As all entries are unconstrained in the worst case all 15 parameters need to be be estimated [76], as the matrix is defined up to scale. In most clinical application the problem can be simplified due to the fact that all X-ray machines are calibrated and its parameters can be retrieved.

4.1.5. Non-linear transformations

In contrast to the previously mentioned global transformations (rigid, affine and projective) non-linear or deformable transformation allow for modeling local transformations. Such models can then be used e.g. to compensate for breathing motion, cardiac motion etc. The field of non-linear registration has become an important research topic in the recent decade [135].

In the case of deformable registration an transformation model is defined on a set of control point, usually placed on a grid. The interpolation of the aligned moving image to the fixed image is defined using basis functions such as Thin-Plate splines, B-Splines, Radial basis functions etc. The amount of parameters is much larger than with the regular, global transformation models. In most cases a gradient based optimization approach is used to estimate the parameters.

In many non-linear registration methods a dense-field parametrization is utilized whereby a deformation field $\vec{u}$ is used as a parametrization to map every voxel in the moving image to the fixed image.

$$\vec{\theta}(\mathcal{M}(x)) = \phi(\mathcal{M}(x + \vec{u}(x)) \qquad u : \Omega \rightarrow \Omega \tag{4.10}$$

In most instances an energy minimization formulation can be derived from the previous formula:

$$S\left(\mathcal{F}, \vec{\theta}(\mathcal{M}) + \alpha R(\vec{u})\right) \tag{4.11}$$

Hereby S depicts the dissimilarity between the fixed image $\mathcal{F}$ and the transformed moving image $\mathcal{M}$. R depicts the regularization factor on the deformation field $\vec{u}$. It should penalize unlikely physical deformations of the tissue.

4.1.6. Similarity Measures

Usually there is no closed form solution in order to optimize the parameters of the transformation model [74]. Thus in most scenarios an iterative approach is used. In each iteration a similarity measure S is used to compare the two images in respect with the current parameters $\vec{\theta}$ of the transformation model. An ideal similarity measure should have its global maximum when the fixed $\mathcal{F}$ and moving image $\mathcal{M}$ are best aligned. In addition it should smoothly approach the maximum while optimizing the transformation parameters and should be insensitive to different types of noises and robust against smaller outliers.

In respect to the similarity measure S the general registration formulation can be reformulated as

$$\vec{\theta}^* = \arg\max_{\vec{\theta}} S\left(\{f_y, m_y\}\right) \qquad f_y = \mathcal{F}(\vec{x}_y); m_y = \mathcal{M}(\vec{\theta}(\vec{x}_y)) \tag{4.12}$$

where f_y defines the intensity in the fixed image $\mathcal{F}$ and m_k the intensity in the moving image $\mathcal{M}$ at the transformed location $\vec{\theta}(\vec{x}_y)$.

The maximization of the similarity measure can also been seen as the minimization of an error metric between the alignment of the two images which can usually be expressed as the inverse or negative form of the similarity measure. Which specific similarity measure should be used depends largely on the problem set and the underlying images which need to be aligned. In the following section the most popular used similarity measures will be explained.

Sum of Squared Differences (SSD)

The sum of squared differences (Sum of squared differences (SSD)) is an error measurement which sums the squared differences of the image intensities between the two images. The error is minimal if the images are perfectly aligned and having similar intensities. Thus only modalities where the intensities between anatomies correspond in the two images can be used in conjunction with SSD. In most cases this applies to images from the same same modality [75] or similar modalities, such as CT and 3D C-arm CT [121, 113, 202].

It can be formulated as follows

$$SSD = \frac{1}{N}\sum_{y=1}^{N}(f_y - m_y)^2 \tag{4.13}$$

where N represents the number of points in the region (or the whole image) where the SSD error metric is evaluated. Even within the same modalities outliers can easily let the method fail, such as different acquisition protocols (MRI) or different contrast phases and types (CT). In addition the SSD is very sensitive to any type of noise except stationary Gaussian noise.

A common modification to reduce the sensitivity to outliers is to us the Sum of Absolute Differences (Sum of absolute differences (SAD)) instead of SSD:

$$SAD = \frac{1}{N}\sum_{y=1}^{N}|f_y - m_y| \tag{4.14}$$

The disadvantage of SAD is the non-differentiability at zero which makes it harder to compute derivatives.

Normalized Cross-Correlation (NCC)

Normalized Cross-Correlation (Normalized Cross Correlation (NCC)) is a very common technique used in signal processing to assess the similarity between two signals (waveforms). It computes the amount of linear correlation between the two signals by computing the product of their normalized values (mean is subtracted from the signal), divided by their standard deviation.

$$NCC = \frac{1}{\sigma_i \sigma_j}\sum_{y=1}^{N}\left(f_y - \overline{f}\right)\ \left(m_y - \overline{m}\right) \tag{4.15}$$

$$\text{with } \overline{f} = \frac{1}{N}\sum_{y=1}^{N} f_y \quad \sigma_f = \sqrt{\frac{1}{N}\sum_{y=1}^{N}\left(f_y - \overline{f}\right)^2} \tag{4.16}$$

$$\text{with } \overline{m} = \frac{1}{N}\sum_{y=1}^{N} j_y \quad \sigma_m = \sqrt{\frac{1}{N}\sum_{y=1}^{N}\left(m_y - \overline{m}\right)^2} \tag{4.17}$$

The arithmetic mean is used to compute $\overline{f}$ and $\overline{m}$. Thus the final formula for the NCC is defined as

$$NCC = \frac{\sum_{y=1}^{N}\left(f_y - \overline{f}\right)\ \left(m_y - \overline{m}\right)}{\left(\sqrt{\frac{1}{N}\sum_{y=1}^{N}\left(f_y - \overline{f}\right)^2}\right)\left(\sqrt{\frac{1}{N}\sum_{y=1}^{N}\left(m_y - \overline{m}\right)^2}\right)} \tag{4.18}$$

Normalized cross correlation has be widely used in mono-modal registration tasks [8, 21, 149, 81, 161]. The limitation is that in the common form a linear relationship between the intensities of the two images is assumed.

Mutual Information

The mutual information metric exploits the property that if two images are correctly aligned the amount of information in order to describe their representation is minimal.

$$MI = \sum_{f}\sum_{m} P_{\mathcal{FM}}(f,m)\log\frac{P_{\mathcal{FM}}(f,m)}{P_{\mathcal{F}}(f)P_{\mathcal{M}}(m)} \tag{4.19}$$

whereby $P_{\mathcal{FM}}$ symbolizes the joint probability distribution of the $\mathcal{F}$ and $\mathcal{M}$ image, and $P_{\mathcal{F}}$ and $P_{\mathcal{M}}$ the marginal probability distribution of $\mathcal{F}$ and $\mathcal{M}$ respectively. Shannon

entropy is a common method used as an indication for the amount of information needed to represent an image:

$$H(\mathcal{F}) = -\sum_i P_{\mathcal{F}}(f) \log P_{\mathcal{F}}(f) \tag{4.20}$$

If the image contains only a constant value $H(\mathcal{F})$ would be zero since $p \log p$ equals 0. The maximum would be obtained if all the voxel have different values and thus the maximum amount of randomness.

In the case of two signals (images) a joint representation can be derived using intensity tuples from both images $\{f_y, m_y\}$ and their joint probability. The joint entropy of $\mathcal{F}$ and $\mathcal{M}$ is

$$H(\mathcal{F}, \mathcal{M}) = -\sum_f \sum_m P_{\mathcal{FM}}(f, m) \log P_{\mathcal{FM}}(f, m) \tag{4.21}$$

In most cases the individual joint probability distributions are estimated using histograms. Thus the final Mutual Information (Mutual Information (MI)) metric is computed as follows:

$$MI = H(\mathcal{F}) + H(\mathcal{M}) - H(\mathcal{F}, \mathcal{M}) \tag{4.22}$$

The value of the Mutual information is in the range $[0 \dots E_{max}]$, where E_{max} is the maximum entropy of either of the two images. Normalizing the values $[0 \dots 1]$ can be done using the following formula:

$$MI_{normalized} = \frac{2MI}{H(\mathcal{F}) + H(\mathcal{M})} = 2 - \frac{2MI}{H(\mathcal{F}) + H(\mathcal{M})} \tag{4.23}$$

Mutual information is well suited for multi-modal registration. Especially in the application area of registering MRI and CT many commercially solutions are available [132, 123, 98].

4.1.7. Optimization Methods

The goal of the optimization of the optimization algorithm is to find the best parameter vector of the transformation model $\vec{\theta}$ that minimizes the value of a certain cost function F:

$$\vec{\theta}^* = \arg\min_{\vec{\theta}} F(\vec{\theta}) \tag{4.24}$$

As mentioned before usually no closed solution is available to determine the optimal transformation parameters $\vec{\theta}$ and a specific iterative scheme is applied to search within the parameter space. It terminates based on a specific abortion criteria, e.g. if the cost function does not change over a specific number of iterations or falls below a certain limit. For image registration, as the computation of the similarity measure is expensive (as it has to be evaluated on the whole image), it is important to keep the number of iterations low and enable fast convergence. In addition as the similarity measure depends on the image con-

tent, it is usually non-linear. In the following we will describe three common subgroups of optimization methods used for registration: 1) Non-gradient based, 2) Gradient based and 3) Quasi-global methods.

4.1.8. Approximate optimization strategies

Hill Climbing is one of the simplest optimization methods. Based on the current estimate of the parameters the cost function is evaluated for each parameter in a local neighborhood. The best estimate is then chosen as the next parameter during the next iteration. Thus the number of evaluations that have to be evaluated is $2N$, where N is the number of parameters (for a rigid transformation in 3D N would be 6). The algorithm terminates if no better estimate can be obtained from the neighbors of the current parameter estimate. In addition it is important to select an appropriate parameter scaling, such that the effect of individual parameters on the cost function are in the same order of magnitude. Several variants of this methods have been successfully applied for the image registration problem [199, 174, 141].

Simplex Method is a simple and popular algorithm in linear programming [152]. In a 2 dimensional parameter space the simplex starts with 3 evaluations of the system response (of the cost function) obtained with 3 different parameter settings. These 3 observations correspond to the vertices of a triangle constituting the 1st simplex. In the 3-dimensional space 4 initial observations are required defining a tetrahedral body. From the evaluation of the response of each observation the position of next point to be evaluated (parameter values within next iteration step) is indicated by either reflection, expansion, or contraction operations.

Powell algorithm starts at a given position in the parameter space, and minimizes the cost function successively along certain directions. Therefore the problem is split up in two parts: Finding the best directions in n-dimensional space, and doing efficient line minimization on a new cost function with only one parameter.

4.1.9. Gradient based optimization strategies

If the gradient vector of F is available, it can be used as base for determining the next steps, as done in the Gradient Descent or Conjugate Gradient methods. Based on a current estimate of the parameters $\vec{\theta}_k$ the update can be calculated as follows:

$$\vec{\theta}_{k+1} = \lambda \frac{\partial F(\vec{\theta}_k)}{\vec{\theta}_k} \tag{4.25}$$

Hereby λ is defining the learning rate. Many improvements have been proposed, especially with the motive to achieve faster convergence. Such methods are called Quasi-Newton approaches where an estimate of the Hessian matrix of F is computed in order to determine the update of $\vec{\theta}$.

4.1.10. Quasi-global methods

Both non-gradient and gradient approaches have the tendency to be trapped in local minima. In order to overcome this limitation Quasi-global approaches repeatedly start the optimization from different starting points in the parameter space, eventually choosing the best result as the optimum parameters $\vec{\theta}^*$. In order to avoid local optima some methods add noise to the cost function value, either purposely, or as byproduct of an randomized cost function approximation. This results in a stochastic optimization approach.

4.2. Model-Based Registration using the Trachea Bifurcation Model

In this chapter we introduce a novel method to estimate the transformation parameters $\vec{\theta}$ between the fixed image $\mathcal{F}$ and $\mathcal{M}$ by using geometric models. If the geometric models, which either contain the anatomy of interest or anatomies which correlate with it, can be robustly estimated from both images a transformation $\vec{\theta}$ can be derived based on the models and thus the two modalities can be brought into correspondence. As our goal is to align a pre-operative CT and the interventional 3D C-arm CT for cardiovascular minimally invasive procedures, we selected the trachea bifurcation model as our anchor anatomy. The trachea bifurcation model is in close proximity to the heart, is seen on both modalities without adding contrast agent and its location correlates well with the location of the heart [29, 121, 189]. Thus based on the estimated trachea bifurcation models the modalities can be aligned.

As we use the trachea bifurcation model as a surrogate anatomy to map the aortic valve from the pre-operative CT to the intra-operative 3D C-arm CT, the modeling and automated estimation of the anatomy is a critical part of our framework.

4.2.1. Trachea Modeling and Estimation

The global position of the trachea bifurcation model is parametrized with a similarity transformation in the three-dimensional Cartesian space, illustrated as a bounding box in Fig. 4.3.

$$\Theta = \{(c_x, c_y, c_z), (\phi_x, \phi_y, \phi_z), (s_x, s_y, s_z)\} \tag{4.26}$$

where (c_x, c_y, c_z), (ϕ_x, ϕ_y, ϕ_z), (s_x, s_y, s_z) are representing the position, orientation (as Euler angles) and scale parameters. The parameters are estimated within the marginal space framework. A probabilistic boosting tree classifier is trained based on Haar features for position and steerable features for orientation and scale, as described in [205].

The next modeling layer consists of four landmarks which are defining key anatomical properties of the trachea bifurcation model: The trachea airway bifurcation point t_B, the trachea lower-left airway branching point t_{LL}, the trachea lower-right airway branching point t_{LR} and the trachea upper center airway point t_T (see Figure 4.3). Using the prob-

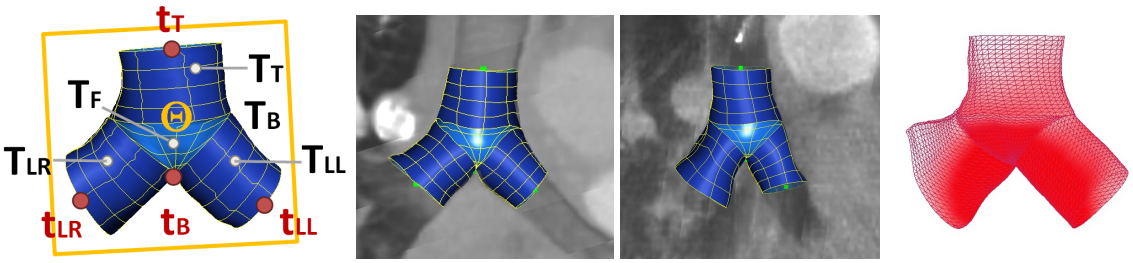


Figure 4.3.: The trachea bifurcation model. Left: Model showing the bounding box Θ, anatomical landmarks: trachea airway bifurcation point t_B, trachea lower-left airway branching point t_{LL}, trachea lower-right airway branching point t_{LR}, trachea upper center airway point t_T and surfaces: trachea upper center airway model T_T, trachea lower-left airway model T_{LL} and trachea lower-right T_{LR} airway model, trachea bifurcation face T_F and back T_B. Trachea bifurcation model shown in CT and 3D C-arm CT (center). The right image is showing the trachea bifurcation model with weights assigned to each vertex for pre-operative to intra-operative mapping A: white - low significance, red - high significance.

abilistic boosting tree classifier (PBT) and Haar features the location of each landmark is learned independently.

The full geometry of the trachea bifurcation is modeled using five surface meshes constructed along rectangular grids of vertices: The underlying grid of each surface mesh is spanned along two physiologically aligned parametric directions, $\vec{u}$ and $\vec{v}$. The trachea upper center airway model T_T, the trachea lower-left airway model T_{LL} and the trachea lower-right T_{LR} airway model are represented as a tubular grid with circumferential $\vec{u}$ and ascending $\vec{v}$ directions including 36×20 vertices. The trachea bifurcation face T_F and back T_B model are represented as paraboloids on a grid of 18×18 vertices. Every corner vertex of the face (T_F) and back model (T_B) has its corresponding vertex on the upper T_T, lower-left T_{LL} and lower-right T_{LR} surface model. The shape model is estimated using the non-rigid Marginal Space Learning framework. Hereby the search space is defined by the first three modes (c_1, c_2, c_3) computed from the statistical shape model. For each mode a boosting classifier is trained using steerable features to distinguish correct hypothesis [205] (see 3.2.3).

4.2.2. Fusion

Our goal is to estimate the rigid transformation $\vec{\theta}$ consisting of a translation $\vec{t}_{\vec{\theta}}$)and a rotation $\vec{\phi}_{\vec{\theta}}$ between the pre-operative CT $\mathcal{F}$to the intra-operative 3D C-arm CT $\mathcal{M}$:

$$\vec{\theta} = (\vec{\phi}_{\vec{\theta}}, \vec{t}_{\vec{\theta}}) \quad ; \qquad \mathcal{F} = \vec{\theta}(\mathcal{M}) \tag{4.27}$$

Based on the detected trachea bifurcation meshes T_1 and T_2 a least squares solution is employed to estimate $\vec{t}_{\vec{\theta}}$ and $\vec{\phi}_{\vec{\theta}}$ [182]. In order to minimize the mapping error in regard to the anatomy of interest, the aortic valve, we estimate a weighting factor for each mesh point.

Based on the aortic valve hinges (3 points) and the aortic valve commissures (3 points) a ground-truth mapping $\vec{\theta}_{GT}$ is estimated. Every intra-operative model of the trachea bifurcation T_2 is transformed to the pre-operative using: $T_{21} = \vec{\theta}_{GT}\ T_2$ and the variance of the point-wise distance $|T_{21} - T_1|$ is computed. The weighting factor $w(i)$ for each mesh point is computed as:

$$w(i) = k\ \frac{1}{M} \sum_{j=1}^{M} |T_{21}(j,i) - T_1(j,i)| \qquad i = 1\ \ldots\ N \tag{4.28}$$

where N is the number of mesh points on the trachea bifurcation surface, M the number of pair-wise mesh annotations and k a normalizing factor. The weights w are later used to solve the weighted least squares mapping:

$$e^2(\vec{\phi}_{\vec{\theta}}, \vec{t}_{\vec{\theta}}) = \frac{1}{N} \sum_{i=1}^{N} \left|\left| w(i)^{-1}\ \ (T_2(i) - T_1(j,i)) \right|\right|^2 \tag{4.29}$$

The solution to this problem can be found in [182]. Fig. 4.3 shows the mean trachea bifurcation model with weights w assigned to each vertex.

4.2.3. Experimental Results

In order to validate the mapping $\vec{\theta}$ from pre-operative CT to the intra-operative 3D C-arm CT we used 28 patient pairs (56 volumes) for our quantitative experiments. The aortic valve was annotated in the contrasted 3D C-arm CT. As we used contrasted intra-operative data the aortic valve was visible and could be delineated manually. Additional 20 test data sets, without pairs of pre-operative and intra-operative images were used to evaluate the estimation performance of the trachea bifurcation model in both the pre-operative CT and intra-operative 3D C-arm CT. All ground-truth annotations were obtained by experts manually placing bounding boxes, anatomical landmarks in the pre- and intra-operative images and finally delineating the surface of the trachea bifurcation and the aortic valve model.

Model estimation

As our weighted mapping relies on the automatic detection of the trachea bifurcation model T we first compare the estimation results on CT and 3D C-arm CT independently. Each detector was trained based on the 28 training data and evaluated on 10 unseen volumes. The accuracy on training and test data is shown in Table 4.1 and 4.2. We compared the estimation of the bounding box Θ, the 4 bifurcation landmarks (t_B, t_{LL}, t_{LR}, t_T) and the combined mesh-to-mesh error from all 5 surfaces (T_T, T_{LL}, T_{LR}, T_F, T_B).

Table 4.1.: System precision for the estimation of the trachea bifurcation bounding box Θ in CT and 3D C-arm CT.

	Training		Testing	
	CT	C-arm CT	CT	C-arm CT
Center [mm]	1.90 ± 0.69	1.98 ± 0.80	3.20 ± 1.58	3.12 ± 2.46
Angle [deg]	3.21 ± 1.44	3.07 ± 1.26	8.08 ± 10.26	9.26 ± 7.90
Scale [mm]	3.41 ± 1.70	2.73 ± 1.36	2.65 ± 0.86	3.09 ± 0.73

Table 4.2.: System precision for the estimation of trachea bifurcation landmarks (t_B, t_{LL}, t_{LR}, t_T) and the surface ($T = T_T \cup T_{LL} \cup T_{LR} \cup T_F \cup T_B$) in CT and 3D C-arm CT.

	Training		Testing	
	CT	C-arm CT	CT	C-arm CT
t_B [mm]	1.96 ± 0.63	1.86 ± 0.80	2.68 ± 1.08	3.30 ± 2.22
t_{LL} [mm]	3.54 ± 1.47	3.62 ± 1.29	4.70 ± 1.39	4.14 ± 2.78
t_{LR} [mm]	3.74 ± 2.19	3.70 ± 2.06	5.16 ± 2.95	5.52 ± 2.72
t_T [mm]	3.22 ± 1.40	2.47 ± 1.21	4.30 ± 1.69	4.30 ± 2.01
T [mm]	0.82 ± 0.11	1.18 ± 0.52	0.87 ± 0.14	1.25 ± 0.62

Model fusion

In order to validate the accuracy of our mapping from the pre-operative CT to the intra-operative 3D C-arm CT the aortic valve was detected in CT (see 3) and mapped to the intra-operative modality using our weighted mapping function A. The estimation errors could be assessed in Table 4.3. Figure 4.5 shows examples of fused volumes with the mapped aortic valve model detected in pre-operative CT and mapped to the non-contrasted 3D C-arm CT.

Table 4.3.: System precision for the estimation of aortic valve in 3D C-arm CT using the mapping estimated from the detected surrogate trachea bifurcation models T_1 and T_2 and the weighted transform A.

	Mean	STD	Median
AV Surface [mm]	7.57	3.22	8.22
Angulation [deg]	9.08	7.31	6.20

4.3. Model-Based Sparse Matching Fusion

In this section we propose an updated registration algorithm capable to register high-quality pre-op data and low-quality intra-op data in the context of minimally invasive car-

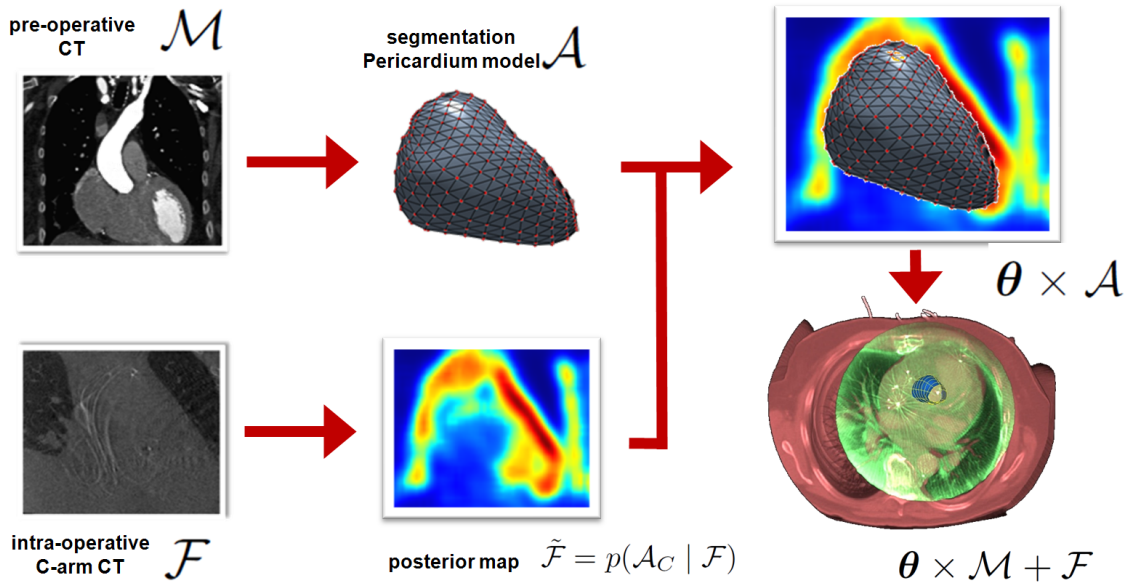


Figure 4.4.: Simplified workflow overview of the model based fusion approach.

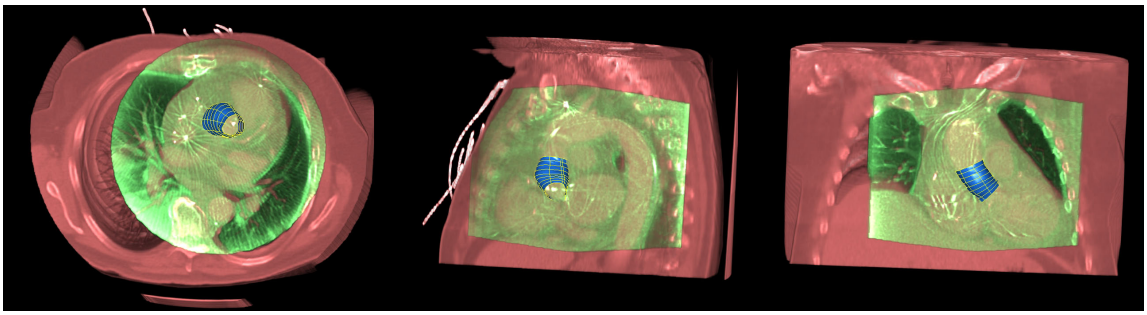


Figure 4.5.: Example of the aligned pre-operative CT image I_1 (red channel) with the intra-operative 3D C-arm CT I_2 (green channel) and the mapped aortic valve model.

diac procedures. Our fully-automatic method registers pre-operative CT (moving image $\mathcal{M}$) with intra-operative C-arm CT volumes (fixed image $\mathcal{F}$), such that a target anatomy, e.g. the aortic valve (AV), is aligned. The process is based on a personalized anchor anatomical model $\mathsf{A}_\mathcal{M}$ of the pericardium extracted from $\mathcal{M}$, and a probability map $\tilde{\mathcal{F}}$ derived from $\mathcal{F}$, see Fig. 4.6. A set of optimal transformation parameters $\vec{\theta}^*$ that align $\mathcal{M}$ with $\mathcal{F}$ is sought. The parameter vector $\vec{\theta} = (\vec{\phi}_{\vec{\theta}}, \vec{t}_{\vec{\theta}})^\top$ represents a rigid transformation in 3D space with $\vec{\phi}_{\vec{\theta}} = (\phi_\mathrm{x}, \phi_\mathrm{y}, \phi_\mathrm{z})$ denoting the Euler angles, and $\vec{t}_{\vec{\theta}} = (t_\mathrm{x}, t_\mathrm{y}, t_\mathrm{z})$ the translation along the axes of the coordinate system (x, y and z).

4.3.1. Pericardium Segmentation

Recently, Zheng et al. [207] presented an efficient and robust method for automatic heart isolation in CT scans [207] by segmenting the pericardium, a double-layered membrane surrounding the heart. Their technique consists of three main steps. First, the pose and scale of the heart is estimated using marginal space learning (MSL) [205]. Second, a mean shape generated from a large number of annotated pericardium meshes is aligned to the estimated pose and scale. In a third step, the parameters are refined within the framework of statistical shape models (SSM) [78] using a boundary detector based on the probabilistic boosting tree (PBT) [179], followed by additional postprocessing. We use this method to segment the patient-specific anchor anatomy $\mathsf{A}_{\mathcal{M}}$ (pericardium mesh) in $\mathcal{M}$. Figure 4.7 shows intersections of a mesh extracted from a CT scan. The mesh is a closed triangulated surface consisting of 514 vertices and 1024 triangles. Since $\mathsf{A}_{\mathcal{M}}$ is independent from intra-operative information, the model segmentation can be performed prior to the intervention.

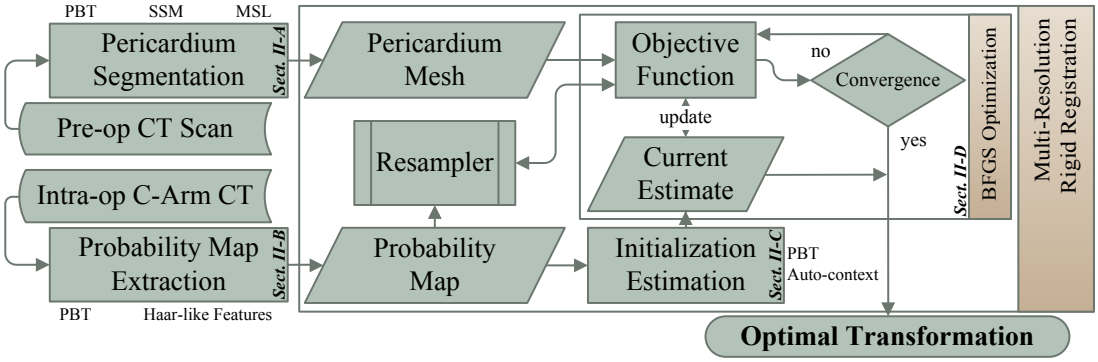


Figure 4.6.: Fusion workflow overview.

4.3.2. Probability Map Extraction

The second data structure is a posterior probability map $\tilde{\mathcal{F}} = p(\mathsf{A}_{\mathcal{F}} \mid \mathcal{F})$ created from $\mathcal{F}$ by evaluating a discriminative classifier on each voxel. We use a PBT classifier that was trained to robustly delineate pericardium boundary regions $\mathsf{A}_{\mathcal{F}}$ in C-arm CT images utilizing 3D Haar-like features [190] to ensure robustness and computational efficiency. In collaboration with medical experts, we created a database $\mathrm{DB} = \{(\mathcal{V}_i, \mathcal{P}_i) \mid i = 1 \ldots n_{\mathrm{DB}}\}$ of ground-truth pericardium meshes $\mathcal{P}_i$ on a set of $n_{\mathrm{DB}} = 393$ interventionally acquired C-arm CT volumes $\mathcal{V}_i$. For training, positive samples were generated with regard to the position of the voxels corresponding to the vertices in the ground-truth annotations. The negative samples for each tuple in the database $(\mathcal{V}_i, \mathcal{P}_i) \in \mathrm{DB}$ are based on randomly selected voxels $\vec{x} \in \mathcal{V}_i$, where the distance of $\vec{x}$ to all points in $\mathcal{P}_i$ exceeds a certain threshold. Figure 4.8 shows an exemplary probability map overlaid on axial volume slices.

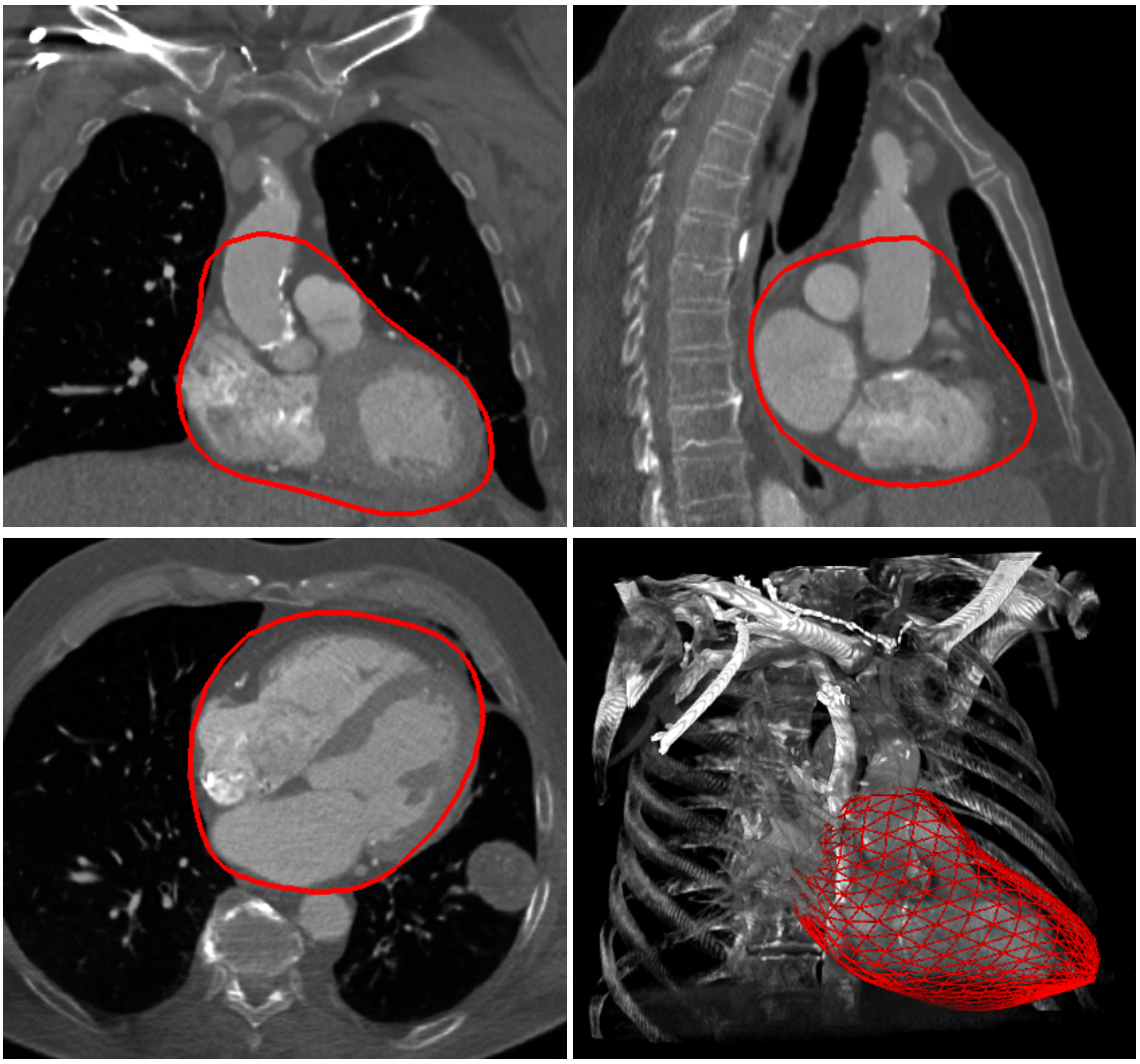


Figure 4.7.: Slices of a volumetric CT scan of the human torso ($\mathcal{M}$) overlaid by the automatically segmented pericardium mesh $\mathsf{A}_{\mathcal{M}}$, the intersection of $\mathsf{A}_{\mathcal{M}}$ with the plane corresponding to the visualized slice is shown in red color. Lower right: 3D rendered CT volume with 3D anatomical overlay ($\mathsf{A}_{\mathcal{M}}$).

4.3.3. Initialization Estimation

One of the major drawbacks in numerical optimization is the need for a reliable initial estimate, i.e. a point in the area of attraction of the global optimum, since without, the method is prone to converge into a local optimum. In order to find such a stable initialization $\vec{\theta}_0 = (\vec{\phi}_{\vec{\theta}_0}, \vec{t}_{\vec{\theta}_0})^\top$, our method recovers the offset $\vec{t}_{\vec{\theta}_0}$ between $\mathcal{M}$ and $\mathcal{F}$. We neglect the rotational error ($\vec{\phi}_{\vec{\theta}_0} = \vec{0}$), since it is rather small between the CT and the C-arm CT scan due to the acquisition protocols being similar as the patients adopt almost identical positions (lying on their back on the table) for both scans.

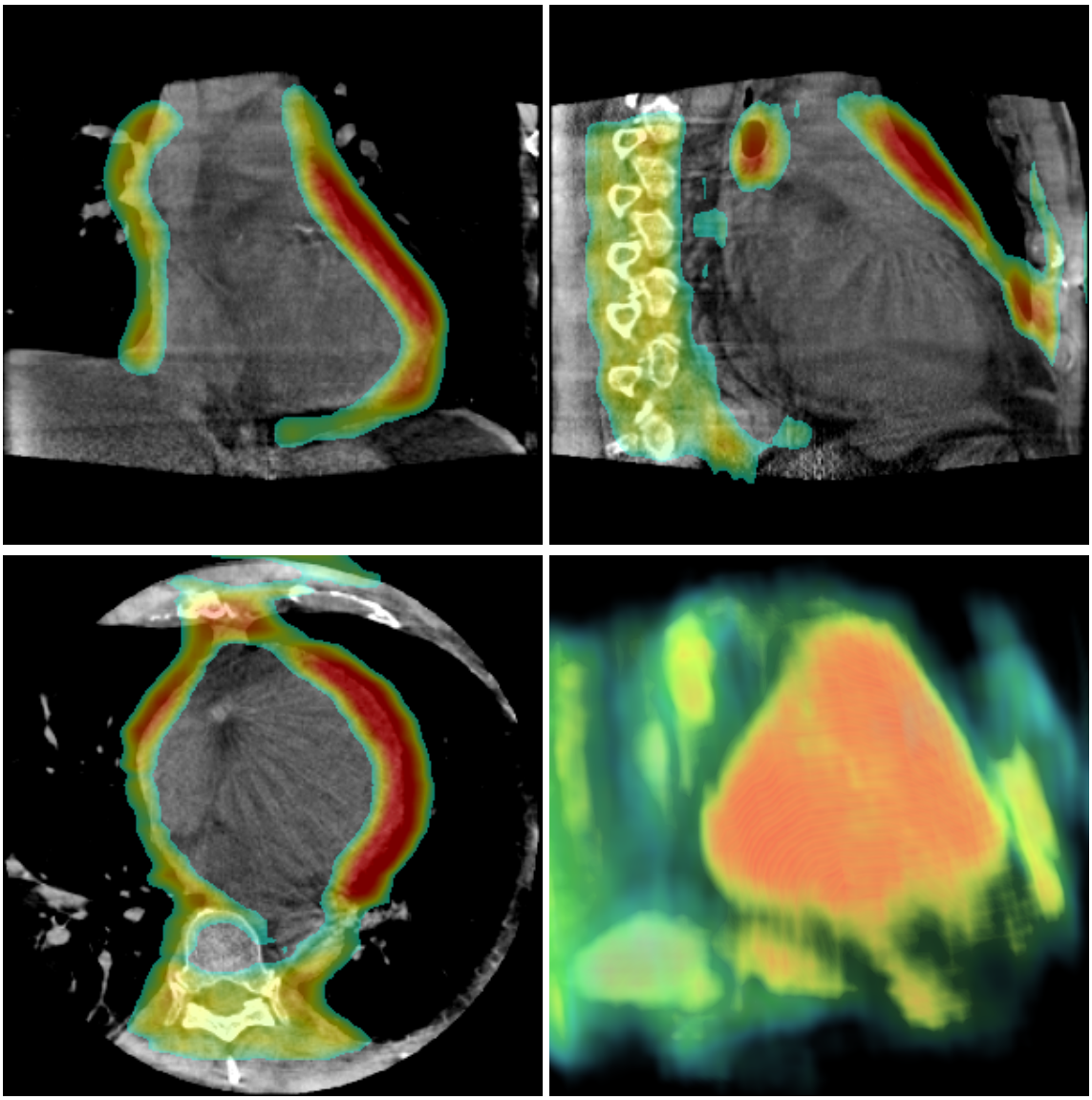


Figure 4.8.: Slices of the original C-arm CT volume ($\mathcal{F}$) overlaid by the PBT-based probability map (thresholded), red color indicates high probability, blue colored and transparent regions are rather unlikely to contain the pericardium boundary. Lower right: Tilted frontal 3D rendering of the probability map.

Our solution is based on a concept from computer vision known as object localization, which we formulate as a classification problem. The object is a single point, the center of the pericardium in the C-arm CT scan, and we aim at locating its position. Therefore, we trained a PBT classifier with Haar-features. In the detection phase, we evaluate the classifier on each voxel and choose the voxel with the maximum probability. Robust detections were achieved by following the idea of Auto-context [180], where a classifier is trained on

the output of another classifier. The probabilistic information from $\tilde{\mathcal{F}}$ is utilized as the input for training and detection, instead of the intensities in $\mathcal{F}$ directly, since the probability maps look similar for both contrasted and non-contrasted images, because the classifier for $\tilde{\mathcal{F}}$ was trained on both types of volumes. This means that the framework is not confused by large magnitudes in intensity gradients, which appear especially in contrasted images at, e.g., the boundaries of the left ventricle or the aorta. Thus, the method can be used with or without contrast agent injected, constituting a major benefit of our work.

For each probability map, a positive training sample was generated for the voxel in $\tilde{\mathcal{F}}$ that is closest to the center of the corresponding annotation $\mathcal{P} \in \mathrm{DB}$ (Sect. 4.3.2), resulting in n_{DB} representatives of the positive class. Negative samples correspond to randomly distributed voxels exceeding a certain distance threshold to the true center.

To summarize, the input to the object localizer is the probability map $\tilde{\mathcal{F}}$ extracted from $\mathcal{F}$, and the output is a pericardium center hypothesis, i.e. the estimated position of the pericardium center in $\mathcal{F}$, say $\vec{c}_{\mathsf{A}_\mathcal{F}}$. Let $\vec{c}_{\mathsf{A}_\mathcal{M}}$ be the center of $\mathsf{A}_\mathcal{M}$. $\vec{\theta}_0$ is then calculated as

$$\vec{\theta}_0 = (\vec{\phi}_{\vec{\theta}_0}, -\vec{c}_{\mathsf{A}_\mathcal{M}} + \vec{c}_{\mathsf{A}_\mathcal{F}})^\top \quad . \tag{4.30}$$

4.3.4. Optimization Strategy

Given a pericardium mesh $\mathsf{A}_\mathcal{M}$, a probability map $\tilde{\mathcal{F}}$ and a starting point $\vec{\theta}_0$, the goal is to refine $\vec{\theta}$ to yield an optimal rigid transformation $\vec{\theta}^*$, such that the anchor anatomy (pericardium), and thus also the target anatomy (aortic valve), is aligned in both images. This process is incorporated into a numerical quasi-Newton minimization framework utilizing the update rule of Broyden, Fletcher, Goldfarb and Shanno (BFGS) [116] for improved convergence behavior and efficiency. We utilize the concept of multi-resolution optimization by optimizing in a coarse-to-fine manner on various granularity levels of $\tilde{\mathcal{F}}$.

Objective Function

Our method iteratively refines a transformation $\vec{\theta}$, which eventually aligns the pericardium in both images, and thus the aortic valve. The optimization problem is defined as

$$\vec{\theta}^* = \underset{\vec{\theta}}{\operatorname{argmin}}\, f(\vec{\theta} \mid \mathsf{A}_\mathcal{M}, \tilde{\mathcal{F}}) \quad . \tag{4.31}$$

The objective function f depends on the mesh $\mathsf{A}_\mathcal{M}$ and the probability map $\tilde{\mathcal{F}}$:

$$f(\vec{\theta} \mid \mathsf{A}_\mathcal{M}, \tilde{\mathcal{F}}) = \frac{\sum_{\vec{p} \in \mathsf{A}_\mathcal{M}} \mathrm{J}(\vec{\theta}(\vec{p}), \tilde{\mathcal{F}}) \cdot \psi(\vec{\theta}(\vec{p}), \tilde{\mathcal{F}})}{\sum_{\vec{p} \in \mathsf{A}_\mathcal{M}} \mathrm{J}(\vec{\theta}(\vec{p}), \tilde{\mathcal{F}})} \quad , \tag{4.32}$$

where $\vec{p}$ denotes a vertex in $\mathsf{A}_\mathcal{M}$ and $\vec{p'} = \vec{\theta}(\vec{p})$ is that vertex transformed w.r.t. $\vec{\theta}$. The indicator function $\mathrm{J}(\vec{p'}, \tilde{\mathcal{F}})$ evaluates to 1, if $\vec{p'}$ is inside the physical boundaries of the volume $\tilde{\mathcal{F}}$, otherwise it returns 0. $\psi(\vec{p'}, \tilde{\mathcal{F}})$ returns a value that is inversely proportional to the probabilistic prediction at the voxel $\vec{x} \in \tilde{\mathcal{F}}$ where $\vec{p'}$ is located.

Gradient Computation

In steepest-descent based minimization, the gradient ∇f is exploited to obtain the descent direction in each iteration. Furthermore, the BFGS method relies on the gradient in order to estimate an approximation of the inverse of the Hessian. Unfortunately, f is highly complex and therefore does not allow for analytical derivations. Hence, we approximate the gradient $\tilde{\nabla} f \approx \nabla f(\vec{\theta} \mid \mathsf{A}_{\mathcal{M}}, \tilde{\mathcal{F}})$ component-wise using the concept of finite differences:

$$\tilde{\nabla} f_i = \frac{1}{\|\delta\|} \left(f(\vec{\theta} + \vec{\delta}^i \mid \mathsf{A}_{\mathcal{M}}, \tilde{\mathcal{F}}) - f(\vec{\theta} \mid \mathsf{A}_{\mathcal{M}}, \tilde{\mathcal{F}}) \right), \tag{4.33}$$

where $\tilde{\nabla} f_i$ denotes the i^{th} component of $\tilde{\nabla} f$. $\vec{\delta}^i$ is a 6D offset vector where all components are zero, except for the i^{th} component $\vec{\delta}^i_i$, which is set to a particular step size. Despite its asymmetric computation scheme, the gradient is sufficiently stable in this application. For the translational components, $\vec{\delta}^i_i$ equals half of the resolution of $\tilde{\mathcal{F}}$. This choice asserts that (4.33) does not evaluate to zero, since a large portion of the points in $\mathsf{A}_{\mathcal{M}}$ transformed w.r.t. $\vec{\theta}$ will correspond to a different voxel $\vec{x} \in \tilde{\mathcal{F}}$ than their corresponding points transformed w.r.t. $\vec{\theta} + \vec{\delta}^i$. Regarding the rotational components, we experimentally determined that a spacing proportional to the resolution of $\tilde{\mathcal{F}}$ (e.g. $\vec{\delta}^i_i = 1^\circ$ when resolution is 1 mm) works properly. Finally, we scale the individual gradient components to equilibrate inconsistencies regarding the magnitude of translation and rotation.

While computing the translational gradient components is straightforward, rotation in 3D poses a major problem due to its inherent non-linearity and co-dependencies. We address these issues by utilizing a linearization of rotation matrices $\vec{R}$ using a first order approximation $\vec{R}'$ as proposed by Mitra et al. [134]. With homogeneous coordinates, rigid transformation turns into a linear problem and thus can be represented by a matrix-vector multiplication with a matrix $\vec{M}_{\vec{\theta}} = (\vec{R}_{\vec{\theta}} \mid \vec{t}_{\vec{\theta}})$, which concatenates rotation and translation into one matrix and represents $\vec{\theta}$. Let $\vec{R}_{\vec{\theta}}$ be the Euler 3D rotation matrix

$$\vec{R}_{\vec{\theta}} = \begin{pmatrix} \cos\phi_{\mathrm{z}} & -\sin\phi_{\mathrm{z}} & 0 \\ \sin\phi_{\mathrm{z}} & \cos\phi_{\mathrm{z}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\phi_{\mathrm{y}} & 0 & \sin\phi_{\mathrm{y}} \\ 0 & 1 & 0 \\ -\sin\phi_{\mathrm{y}} & 0 & \cos\phi_{\mathrm{y}} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi_{\mathrm{x}} & -\sin\phi_{\mathrm{x}} \\ 0 & \sin\phi_{\mathrm{x}} & \cos\phi_{\mathrm{x}} \end{pmatrix}. \tag{4.34}$$

Its first-order approximation is given by

$$\vec{R}'_{\vec{\theta}} = \begin{pmatrix} 1 & -\phi_{\mathrm{z}} & \phi_{\mathrm{y}} \\ \phi_{\mathrm{z}} & 1 & -\phi_{\mathrm{x}} \\ -\phi_{\mathrm{y}} & \phi_{\mathrm{x}} & 1 \end{pmatrix} \approx \vec{R}_{\vec{\theta}} \ . \tag{4.35}$$

It is important to mention that $\vec{R}'_{\vec{\theta}} \approx \vec{R}_{\vec{\theta}}$ only holds under small motion, i.e. when

$\|\vec{\phi}\|_2 \to 0$. Hence, we cannot use $\vec{R}'_{\vec{\theta}}$ for large angles without introducing errors. Therefore, we compute the rotational components of $\tilde{\nabla}$ using a composite transformation. First, a point $\vec{p} \in \mathsf{A}_{\mathcal{M}}$ is transformed w.r.t. the current estimate $\vec{\theta}$ using the exact Euler-angle representation to generate an intermediate point $\vec{p''}$. Second, $\vec{p''}$ is rotated according to the minor rotation $\vec{\delta^i}$ to yield $\vec{p'}$ by making use of the linearization $\vec{R}'_{\vec{\theta}}$ from (4.35). Thus, we get

$$\vec{p'} = \vec{\delta^i}(\vec{p''}) = \vec{\delta^i}(\vec{\theta}(\vec{p})) \ . \tag{4.36}$$

$\vec{p'}$ constitutes the first argument for the functions J and ψ in (4.32) when computing the approximate rotational gradient components of $\tilde{\nabla}$ using (4.33).

Multiple Resolution Approach

To compensate for potential initial coarse misalignment, we exploit the concept of multi-resolution optimization. The optimizer's area of attraction is synthetically augmented by using low-resolution probability maps at first, since gradient computation as in (4.33) is based on offset vectors $\vec{\delta^i}$ whose component values are proportional to the resolution of the volume. First, after convergence of a low-resolution optimizer using a low-resolution probability map $\tilde{\mathcal{F}}|_{\text{coarse}}$, a rough registration $\vec{\theta^*}|_{\text{coarse}}$ is determined. The next optimization step is then performed using a resampled, finer representation of the probability map $\tilde{\mathcal{F}}|_{\text{fine}}$. The optimization w.r.t. $\tilde{\mathcal{F}}|_{\text{fine}}$ starts at $\vec{\theta}|_{\text{coarse}}$ and eventually yields $\vec{\theta^*}|_{\text{fine}}$. This process is repeated until a certain target resolution $\tilde{\mathcal{F}}|_{\text{finest}}$ is reached. The resulting $\vec{\theta^*} = \vec{\theta^*}|_{\text{finest}}$ then constitutes the multi-resolution-optimal set of parameters. We consider three different isotropic scales (4, 2, and 1 mm).

Weighted Sampling

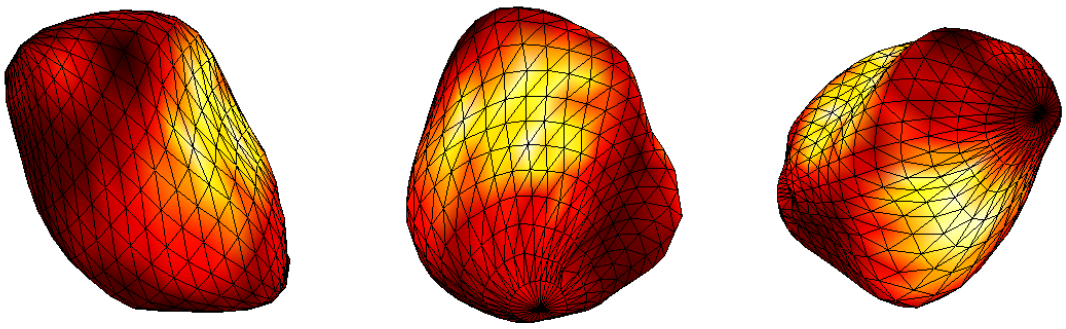


Figure 4.9.: Visualization of the prior sampling regions for $\mathsf{A}_{\mathcal{M}}$, dark/red colors depict regions of sparse sampling, while bright/yellow regions are sampled more densely during the optimization.

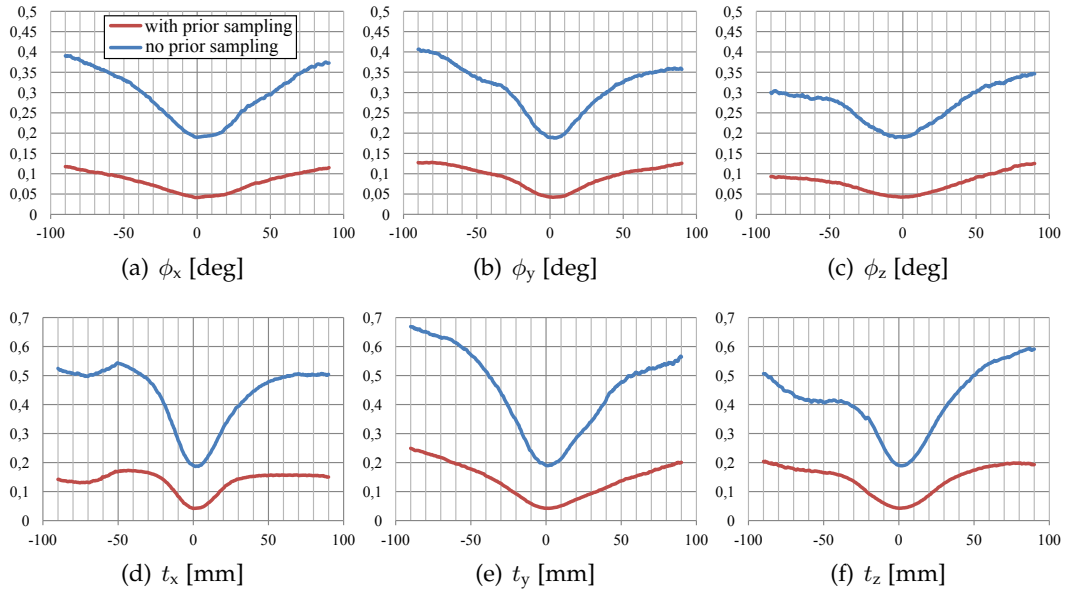


Figure 4.10.: Energy curves of the optimizer, comparing the initial objective function from (4.32) against the more sophisticated objective function leveraging prior knowledge by performing a probabilistic sampling of $\mathsf{A}_{\mathcal{M}}$ from (4.37).

The classifier response (Sect. 4.3.2) is more reliable in some regions of the volumes compared to others. For instance see Fig. 4.8, where the areas close to the left ventricle and right atrium have high responses, while classification near the spine is noisy and the right ventricle region shows low confidence. Robustness and accuracy of our method (Sect. 4.3.8) could be improved significantly by incorporating prior knowledge into the optimization procedure. Therefore, the surface mesh of the anchor anatomy $\mathsf{A}_{\mathcal{M}}$ is subdivided into small regions $r_i \in \mathsf{A}_{\mathcal{M}}$ with $\bigcup_i r_i = \mathsf{A}_{\mathcal{M}}$ and $r_i \cap r_j = \emptyset \; \forall i, j : i \neq j$. Each region r_i is assigned a patient-independent weight w_{r_i} yielding a weight vector $\vec{w} = \{w_{r_i} \mid r_i \in \mathsf{A}_{\mathcal{M}}\}$. $\vec{w}$ increases the influence of those regions that are likely to be located within a high confidence area in $\tilde{\mathcal{F}}$, whereas a region that is noisy or often falsely classified gets penalized. This is achieved by increasing the sampling rate of the pericardium mesh at regions r_i with high weights w_{r_i}, while reducing the sampling rate for low w_{r_i} regions. Gradient magnitude, as well as the distance to the upper and lower boundary of the pericardium were most useful for reliability predictions. Based on these observations and DB from Sect. 4.3.2, we computed values for $\vec{w}$ (cf. Fig. 4.9).

Let $g_{\vec{w}}$ be a probabilistic function that returns three-dimensional points on the surface of a given mesh according to the sampling probabilities defined by $\vec{w}$. To incorporate the prior knowledge into our framework, we extend the objective function f from (4.32) and obtain:

$$f(\vec{\theta}\,|\,\mathsf{A}_{\mathcal{M}}, \tilde{\mathcal{F}}, g_{\vec{w}}) = \frac{\sum_{\vec{p}\in g_{\vec{w}}(\mathsf{A}_{\mathcal{M}})} \mathrm{J}(\vec{\theta}(\vec{p}), \tilde{\mathcal{F}}) \cdot \psi(\vec{\theta}(\vec{p}), \tilde{\mathcal{F}})}{\sum_{\vec{p}\in g_{\vec{w}}(\mathsf{A}_{\mathcal{M}})} \mathrm{J}(\vec{\theta}(\vec{p}), \tilde{\mathcal{F}})} \,. \tag{4.37}$$

In Fig. 4.10, we compare energy curves for the optimizer with standard sampling against energy functions of an optimizer that follows (4.37) with prior sampling for one exemplary dataset. Starting at the estimated optimal point in parameter space, i.e. after registration, the values for the plots are computed by either translating the pericardium in one direction (x, y or z) or rotating the pericardium around the vector defined by its center and one of the coordinate axes and then evaluating the objective function f given the manually misaligned $\mathsf{A}_{\mathcal{M}}$ and $\tilde{\mathcal{F}}$. Prior sampling significantly reduces noise and allows for energy functions that are more smooth and provide large areas of attraction (convex regions around the global optimum) with less local minima. For instance see Fig. 4.10(f), where monotonicity significantly fails at -20 mm in the blue curve, while the red curve (prior sampling enabled) stays nicely monotonous.

4.3.5. Experimental Results

4.3.6. Dataset and Error Measure

Clinical Dataset

We compiled a set of 95 corresponding clinical CT and C-arm CT volumes, each with an isotropic resolution of 1 mm. 25 image pairs are native, while 70 were acquired with contrast agent injected into the aorta. Medical experts annotated the pericardium in each volume, 43 studies include annotations of the aortic valve (AV). The data is organized in a database $\mathrm{DB}_{\text{clinic}} = \{(\mathcal{V}_i, \breve{\mathcal{V}}_i, \mathsf{P}_i, \breve{\mathsf{P}}_i, \mathsf{R}_i, \breve{\mathsf{R}}_i) \mid i = 1 \ldots 95\}$. $\mathcal{V}_i$, P_i and R_i denote i^{th} CT volume, its pericardium and AV annotation, respectively. Analogously, $\breve{\mathcal{V}}_i$, $\breve{\mathsf{P}}_i$ and $\breve{\mathsf{R}}_i$ denote these structures for the C-arm CT acquisition.

Mesh-to-Mesh Error

Since our method is based on geometrical models and the clinical dataset includes ground-truth annotations, quantitative evaluation is possible. Our results are based on a symmetric mesh-to-mesh distance metric ε. Its implementation utilizes the point-to-triangle distance $\varepsilon_{\text{p2t}}(\vec{p}, \triangle)$ between a point $\vec{p}$ and a triangle $\triangle$. Let X, Y be two triangulated meshes with equal amount of vertices. ε is defined as:

$$\varepsilon(\mathsf{X},\mathsf{Y}) = \frac{1}{2}\left(\frac{1}{\|\mathsf{X}\|} \sum_{\vec{p}\in\mathsf{X}} \min_{\triangle\in\mathsf{Y}} \varepsilon_{\text{p2t}}(\vec{p}, \triangle) + \frac{1}{\|\mathsf{Y}\|} \sum_{\vec{p}\in\mathsf{Y}} \min_{\triangle\in\mathsf{X}} \varepsilon_{\text{p2t}}(\vec{p}, \triangle) \right). \tag{4.38}$$

where $\vec{p} \in \mathsf{X}$ defines a vertex on the triangulated mesh X and $\vec{p} \in \mathsf{Y}$ a vertex on the triangulated mesh Y. Note, that (4.38) might underestimate misalignment tangential to the surface. However, it is a fairly *natural* measure closely resembling visual assessment by human experts.

4.3.7. Evaluation on Synthetic Data

In order to evaluate the convergence behavior of the numerical optimization method we used synthetic data. We utilized the approach described by Mitra [134] of Funnels Of Convergions (see figure 4.12) . Hereby we first computed a synthetic probability map $\tilde{\mathcal{F}}_{synthetic} = p(\mathsf{A}_{\mathcal{F}} \mid \mathcal{F})$ based on a pericardium model placed $\mathsf{A}_{\mathcal{F}}$ in the center of the volume $\mathcal{F}$. This is achieved by computing a distance map based on the model $\mathsf{A}_{\mathcal{F}}$ and thresholding within a certain distance $d < n$ mm. The goal is to achieve high responses within a small distance around the pericardium surface boundary. We added both Gaussian and salt and pepper noise to approximate the appearance of the real probability map.

Afterward we rotated the ground-truth pericardium model (around the y-axis) and translated (along the x-z plane) it to generate different initial positions. Figure 4.11 denotes the sampling pattern used to get the initial positions. The rotation angle is sampled at 10 degree intervals, while the maximum radial translation of the pericardium was selected in 10 mm interval (maximum of 30 mm). Please note that the mean initialization error using our position detector trained using Auto-Context features is below 10 mm. We apply our optimization strategy as described in 4.3.4 to align the transformed pericardium model to the synthetic probability map. Hereby we determine the residual error after the optimization procedure has completed as the Mesh-to-Mesh metric (4.3.6) between the ground-truth model and the aligned after the optimization. We determined qualitatively that a Mesh-to-Mesh error smaller than 5 mm is correlating with a successful outcome.

Finally we construct the Funnel Of Convergence Plots for our method (see figure 4.12). Regions in black denote convergence to the correct solution. Our algorithm is found to have a broad, and stable convergence funnel. Initializations within 30 mm and ± 40 degree converge to the correct solution.

4.3.8. Evaluation on Clinical Data

Table 4.4.: AAE statistics [mm] $\varepsilon(\vec{\theta}^*(\mathsf{A}_{\mathcal{M}}), \check{\mathsf{P}})$ after registration (#Studies=95)

Method	Mean	Std	Median	80%	90%	Max	Min
Sparse Matching	**5.48**	1.82	5.22	7.26	7.88	10.40	1.93
Sparse Matching (no Prior Sampling)	**7.20**	3.79	6.63	8.85	10.36	33.47	2.11
Quasi-global Search	**13.06**	12.79	5.86	24.38	34.17	57.56	2.33
Quasi-global Search (masked)	**13.69**	19.82	7.03	16.40	37.56	139.74	2.11
ITK Registration Framework	**20.87**	25.05	8.56	33.12	44.34	125.92	2.57
ITK Registration Framework (masked)	**26.20**	25.53	18.26	45.72	59.55	125.92	3.04

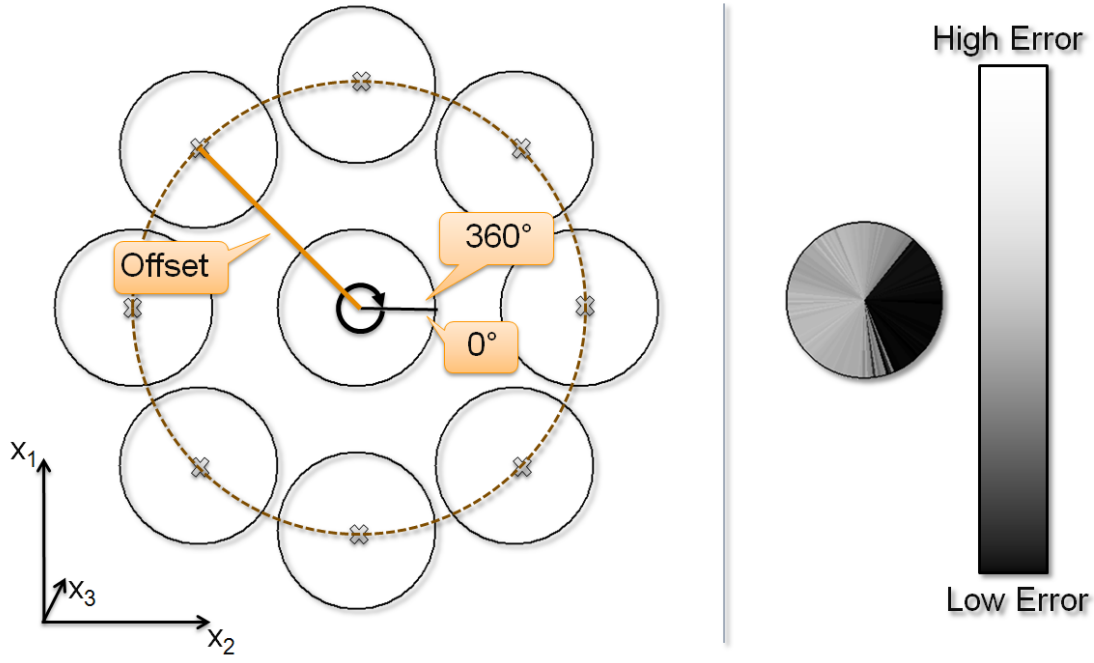


Figure 4.11.: Diagram explaining the funnel convergence plot. Each circle is describing the convergence behavior of our registration method in respect to a certain translational and rotational offset in respect to the ground truth. Black area is symbolizing a small error in respect to the ground truth and white a large error.

Table 4.5.: TAE statistics [mm] $\varepsilon(\vec{\theta}^*(\mathrm{R}), \check{\mathrm{R}})$ after registration (#Studies=43)

Method	Mean	Std	Median	80%	90%	Max	Min
Sparse Matching	**4.67**	1.94	4.22	6.54	7.29	8.83	1.24
Sparse Matching (no Prior Sampling)	**6.14**	3.54	5.56	8.47	9.44	22.35	1.18
Quasi-global Search	**21.76**	22.32	8.49	49.05	60.36	77.67	1.43
Quasi-global Search (masked)	**13.72**	19.73	5.25	18.29	37.50	79.80	1.02
ITK Registration Framework	**34.02**	42.77	9.54	65.21	95.27	161.12	1.21
ITK Registration Framework (masked)	**37.86**	36.59	28.32	70.73	80.66	161.12	1.07

We quantitatively evaluated our fusion approach on $\mathrm{DB}_{\mathrm{clinic}}$ with both, prior sampling (Sect. 4.3.4) enabled and disabled. The first three rows in Table 4.4 show the anchor anatomy alignment error (AAE) statistics, i.e. error statistics resulting from a comparison of the optimally transformed segmented pericardium $\vec{\theta}^*(\mathrm{A}_{\mathcal{M}})$ and the ground-truth annotation in the C-arm CT volume $\check{\mathrm{P}}$. Table 4.5 shows the target anatomy errors (TAE), i.e. a comparison of the transformed CT-based aortic valve $\vec{\theta}^*(\mathrm{R})$ and its C-arm CT based annotation $\check{\mathrm{R}}$. From left to right, the columns contain the name of the analyzed method, followed by error measurements in mm, starting with the mean error and standard deviation, the 50^{th} (median), 80^{th} and 90^{th} percentiles of the errors, as well as the maximum

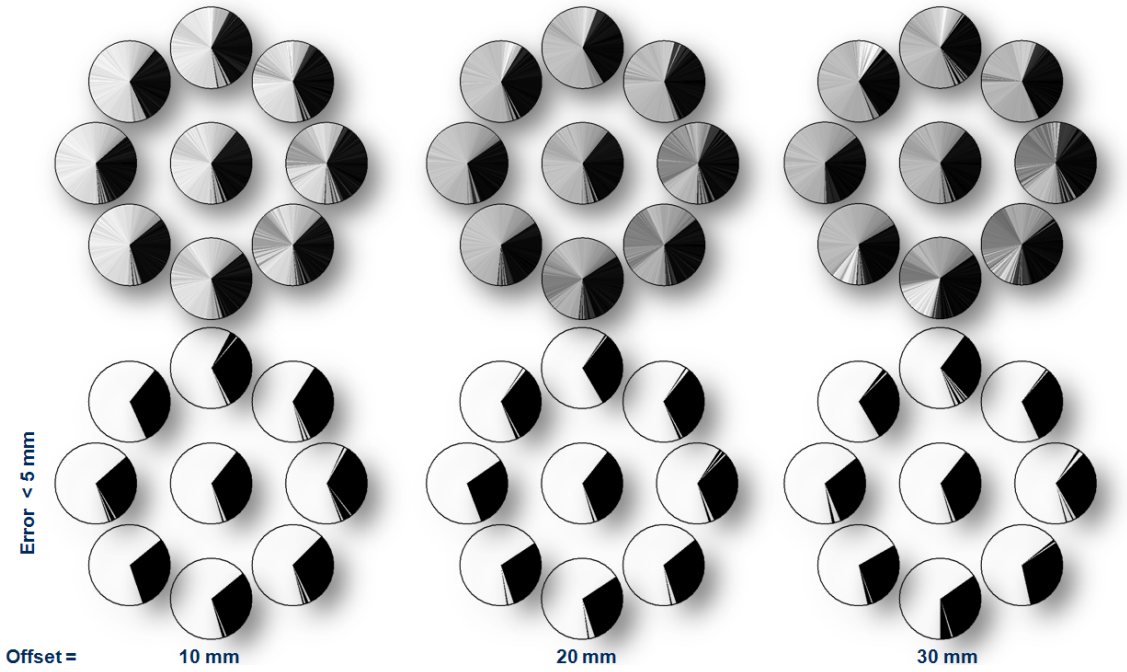


Figure 4.12.: Diagram showing the convergence behavior of our optimization method using the funnel convergence plot on synthetic data. Black area is symbolizing a small error in respect to the ground truth and white a large error. Result shows that we can recover offsets of ± 40 degree and 30 mm robustly.

Table 4.6.: AAE statistics (contrasted only) [mm] $\varepsilon(\vec{\theta}^*(\mathsf{A}_{\mathcal{M}}), \check{\mathsf{R}})$ after registration (#Studies=70)

Method	Mean	Std	Median	80%	90%
Sparse Matching	**5.33**	1.69	5.22	6.58	7.73
Sparse Matching (no Prior Sampling)	**7.30**	4.21	6.69	8.82	10.92
Sparse Matching (Initialization only)	**6.15**	2.16	5.67	7.35	9.75
Quasi-global Search	**14.66**	14.06	6.61	30.42	39.13
Quasi-global Search (masked)	**13.49**	20.15	6.70	16.40	42.88
ITK Registration Framework	**22.59**	28.02	8.41	41.48	55.21
ITK Registration Framework (masked)	**25.70**	24.43	19.26	44.70	59.93

and minimum error. The two most-right columns show the ratio of fail-cases, where we defined a registration accuracy with an AAE (Table 4.4) or TAE (Table 4.5) greater or equal than 2 cm as failed, and an adjusted mean, where we excluded the fail-cases.

With weighted sampling incorporated, a mean AAE of 5.48 ± 1.82 mm measured between the anchor anatomy (pericardium) is achieved. On the 70 contrasted and the 25 non-contrasted volumes, the AAE is 5.33 ± 1.69 mm and 5.91 ± 2.09 mm, respectively. Furthermore, with a mean TAE of 4.67 ± 1.94 mm, the target anatomy (AV) is aligned properly. No fail-cases were observed for both anchor and target anatomy. When the patient-

independent weighting is ignored, the mean errors increase significantly by more than 35% AAE and almost 50% TAE. One reason is that more outliers are generated, which even leads to a fail-case and has a strong influence on the overall errors. In Fig. 4.15(a-d), we depict representative qualitative fusion results from various datasets $\in \mathrm{DB}_{\mathrm{clinic}}$.

4.3.9. Comparison to State-of-the-Art Registration Methods

We quantitatively compared our model-to-image registration to two state-of-the-art image-to-image registration approaches. One might assume that when rather putting a stronger focus on the anchor anatomy (pericardium) than using the entire image, the registration results will improve. Therefore, we implemented an option to mask a region of interest (ROI) in the pre-operative CT image (denoted as "*Name-of-Method* (masked)" in the result tables below). Given such a mask image, the similarity metric only takes into account those voxels in the CT volume, where the corresponding voxel in the mask is set to enabled. The mask image is created by (i) automatic segmentation of the pericardium as described in Sect. 4.3.1, (ii) converting the pericardium mesh into a binary image of the same size, spacing and pose as the CT image, where voxels inside the mesh are set to enabled and all the other voxels are set to disabled, and (iii) dilating the binary image (increasing the size of the ROI) utilizing a spherical structuring element with a radius of 5 mm. Step (iii) ensures that the entire heart as well as a small area around it is included in the ROI. In our experiments, this last step led to significant improvements in robustness, since especially the bordering area between the pericardium and the lungs provides crucial information that can improve the outcome of the similarity metrics (large gradient magnitudes and homogeneous regions).

ITK Registration Framework

The first method utilizes the Insight Segmentation and Registration Toolkit (Insight Segmentation and Registration Toolkit (ITK)), an open source medical imaging library [86]. Since intensities in the CT image and the C-arm CT image do not necessarily correlate, the similarity metric is based on mutual information [130]. We use all voxels in the image and 50 bins for the histogram as proposed by Mattes et al. [130]. Optimal transformation parameters $\vec{\theta}^{*}_{\mathrm{ITK}}$ were obtained by a multi-resolution optimizer for rigid versor transformations. The scales for the components of the versor were adjusted according to [86]. We customized the maximum and minimum step lengths adaptively for each resolution, the maximum number of iterations was set to 200, and we initialized the procedure by aligning the centroid of both volumes. Results obtained running this framework on $\mathrm{DB}_{\mathrm{clinic}}$ are presented in the last two rows of Table 4.4 (anchor anatomy) and Table 4.5 (target anatomy). The method fails for approximately 40% of the studies. However, since this framework is not specifically designed to align the pericardia and no pre-processing was performed, the large number is understandable. Many of the failures occurred for images with significant differences in the size of the field of view between the CT and the C-arm CT acquisition.

The mean error of 20.87 mm yielded by this framework is substantially larger than the error of our method, with a mean of 5.48 mm on the same datasets. A rigid registration is computed in 501.9 ± 427.4 s, i.e. it takes more than eight minutes on average. Please note that no optimizations w.r.t. runtime were implemented.

To increase the focus on aligning the pericardia, we repeated the experiment using an image mask derived from the CT pericardium mesh as described above. In fact, for some images, the alignment could be improved significantly. For instance, for one dataset the AAE is reduced from 17.23 mm to 3.09 mm when utilizing the mask image. However, the number of cases where the results get worse due to the masking prevail (sometimes the AAE increases by an order of magnitude, e.g. from from 3.06 mm to 31.18 mm).

Quasi-global Search

The second method is a quasi-global knowledge-driven registration approach for thoracic-abdominal CT and C-arm CT images designed for image-guided interventions [202]. It has been proposed only recently. Given an intra-operative C-arm CT image, in the first step they create three surrogate 2D images, so-called Anatomy Targeted Projections (ATP). An ATP is a maximum-intensity-like 2D projection, which focuses on a specific anatomy or tissue type (e.g. bone, soft tissue, etc.). This is achieved by projecting the intensity of the voxel with the maximum likelihood of belonging to the targeted tissue type along each projection ray. The use of 2D ATPs instead of the 3D volume can reduce computational costs significantly and thus allows for a large number of similarity metric evaluations within a reasonable time frame. The authors chose an adaption of normalized mutual information (NMI) as similarity metric. Second, multiple starting points in registration parameters space are chosen to approximate a global search, motivated by the assumption that most CT volumes have a larger field of view compared to C-arm CT images. In a third step, the globally optimal candidate is selected by analyzing similarity values and gradients at all starting points. Last, a local multi-resolution optimization is performed, yielding an optimal set of rigid transformation parameters. The authors claim that their method is fast and robust with low target registration and maximum registration errors on 20 datasets.

Quantitative results (Tables 4.4 and 4.5) show that on average, the AAE of the quasi-global search is more than twice as large as the AAE of our method, and the TAE increases from 4.67 mm to 21.76 mm. This indicates that finding a solution to the problem of aligning the pericardium in such heterogeneous images (varying field of view, contrasted and non-contrasted images, etc.) from different modalities is hard and working with image intensities directly might not be sufficient. However, while our learning-based method works particularly well for the application described in this section, it is not well suited for general-purpose registration. In contrast, the quasi-global search is not specialized for a particular application, it is rather an approach designed for handling various registration tasks from a larger problem domain. The observations regarding the quasi-global search when using the pericardium mask are similar as for the ITK registration framework when using the mask. On the one hand, there are datasets, where the alignment improves sig-

nificantly (for an example see Fig. 4.15(e)), on the other hand, in many cases performance decreases. Besides, masking can lead to individual excessive errors as high as 139.74 mm (see Table 4.4), whereas the maximum error of the standard quasi-global search is below 60 mm. Although a smaller number of fail-cases is observed, the details discussed above lead to a slightly higher mean AAE of 13.69 mm for the method with pericardium masking, compared to 13.06 mm without masking. Surprisingly, the TAE (see Table 4.5) in the 43 datasets where annotations of the aortic valve are available decreases from 21.76 mm without masking to 13.72 mm with masking. This is due to the images in this specific subset working better with pericardium masking (AAE = 11.17 ± 11.25 mm) compared to the method without masking (AAE = 16.35 ± 13.26 mm). Furthermore, on our testing machine (see Sect. 4.3.11), we measured a mean runtime of 2.484 ± 1.076 s until a rigid registration is computed.

4.3.10. Inter-user Variability Study

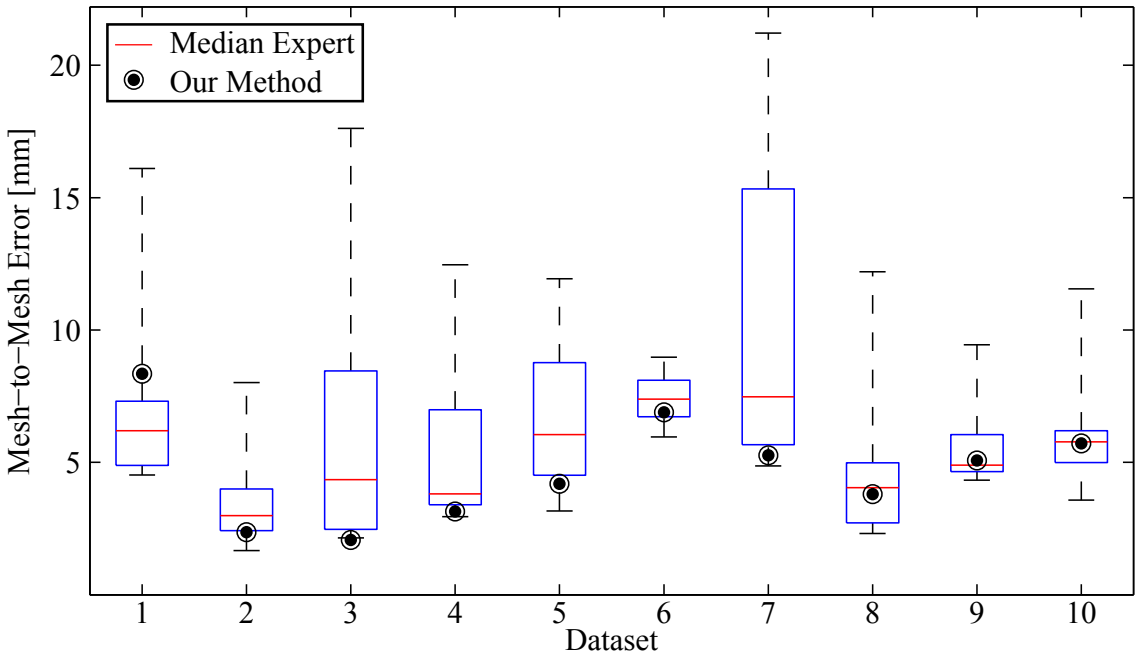


Figure 4.13.: Inter-user variability compared to performance of our method. The edges of the boxes indicate 25$^{\text{th}}$ and 75$^{\text{th}}$ percentiles of the expert errors (AAE).

Ascribing a rational meaning to quantitative results is challenging. In most cases, the true performance of a system would not only be measured in absolute terms but rather relative to the manual performance of experts. Thus, we compared our method to the individual performances of a group of $n_{\text{user}} = 10$ technical experts, who work on such clinical data on a daily basis. The task we assigned to each of the users was to manually align $n_{\text{data}} = 10$ pairs of volumes (a subset of $\text{DB}_{\text{clinic}}$, see Sect. 4.3.6). We provided

them with a software tool that allows for adjusting rigid transformation parameters in a convenient manner and visualizes their progress in real-time.

Let $\vec{\theta}_{ji}$ be the manually estimated transformation parameters of the i^{th} user for the j^{th} pair of volumes. We compare the fit to the ground-truth C-arm CT annotation $\breve{\mathsf{P}}_j$, i.e. we compute the error $\varepsilon(\breve{\mathsf{P}}_j, \vec{\theta}_{ji}(\mathsf{P}_j))$ between the manually annotated fixed C-arm CT pericardium and the moving CT pericardium transformed w.r.t. the user's manual transformation. Results are shown in Fig. 4.13. Our automated method exhibits lower errors than the median user in 80% of all cases and shows high robustness with no outliers. There exists only one pair of volumes, where the automatic fusion is inferior to more than 75% of the users. Moreover, the users' manual fusion time per data pair ranged from two to five minutes, while our method takes less than two seconds on average (see Sect. 4.3.11), which means a speedup of up to 99%. To conclude, with no fail-case and reliable performance, our fully-automatic approach outperforms manual registration in terms of robustness, accuracy and runtime.

4.3.11. Runtime Performance

Our method is designed to be used interventionally, resulting in a need for low computational costs associated with the registration. We conducted a runtime analysis on an off-the-shelf consumer laptop with an Intel® Core™ i7-3720QM CPU @ 2.60 GHz with 4 cores (8 threads) and 8 GB of main memory. Our prototype is implemented in C++ utilizing OpenMP [27] for efficient parallel programming.

The average runtime for the entire process of estimating an optimal rigid transformation $\vec{\theta}^*$ as illustrated in Fig. 4.6 measured on $\text{DB}_{\text{clinic}}$ is 1.562 ± 0.286 seconds. Below, we briefly discuss the runtime behavior of the four major components of the registration framework. For a detailed overview see Fig. 4.14.

We observed that the runtime for the automatic model segmentation (Sect. 4.3.1) from the pre-operative CT volume by Zheng et al. [207] scales linearly with the number of voxels in the volume, ranging from 0.181 s for a scan consisting of $206 \times 206 \times 103$ voxels to 1.70 s for a volume with $497 \times 497 \times 278$ voxels. On average over all CT volumes $\in \text{DB}_{\text{clinic}}$, the mean runtime is 0.451 ± 0.254 s.

The cost for generating a probability map (Sect. 4.3.2) is strongly correlated with the size and resolution of the C-arm CT volume, since the classifier has to be evaluated on each voxel. This is done in parallel on the CPU. Please note that this is a task that can be outsourced to the graphics processing unit (GPU), potentially resulting in a massive gain in performance. For a typically sized volume of $240 \times 240 \times 180\,\text{mm}^3$, probability map generation takes approximately 0.81, 1.94, 5.73 or 39.9 seconds for resolutions of 4, 3, 2 and 1 mm, respectively. In our standard approach, we utilize the 4 mm classifier, i.e. on average, probability map generation accounts for 0.837 ± 0.088 seconds of the overall runtime.

The use of statistical object localization to estimate the initialization $\vec{\theta}_0$ for the optimizer (Sect. 4.3.3) plays only a minor role. We measured a maximum runtime of 24 *milli*seconds.

The last component is the iterative optimization (Sect. 4.3.4) with an average runtime of

0.258 ± 0.042 s, which is mainly influenced by the number of objective function evaluations and by the cost for resampling the probability map (multiple resolutions). The latter needs to be done twice (4 mm→2 mm and 4 mm→1 mm resolution), consuming approximately 0.141 s. The evaluation of f and the approximate gradient computation $\tilde{\nabla}$ is combined in one function g. Averaged over $\mathrm{DB}_{\mathrm{clinic}}$, g was called 144.7 ± 22.1 times (accumulated over all granularity levels). Independent of the current resolution, the mean runtime of one call is 0.81 *milli*seconds. Thus, the average cost of successive calls to g for one registration is 0.117 s.

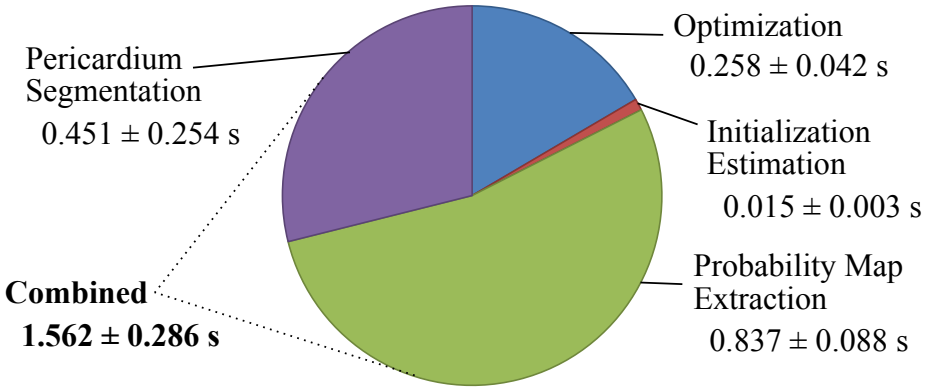


Figure 4.14.: Relative / absolute runtime of the four major framework components.

4.4. Extensions

4.4.1. Extensions

In some clinical setting the acquisition protocol of the 3D C-arm CT is hard to achieve as it requires a spin of 200 degree in a crowded hybrid operating room. Thus a robust fusion algorithm (as seen in the previous section) which can work with limited angle reconstructed 3D C-arm CT is desired. In the next section we propose an extensions of our algorithm to achieve this goal. We demonstrate how our method can be extended to cope with volumes which are reconstructed using a 90 degree sweep (instead of a 200 degree).

4.5. Fusion with Limited-angle Tomosynthesis 3D C-arm CT

Prior weights can be used in order to facilitate registration with an intra-operative Tomosynthesis volume, a volume that was reconstructed using only a constrained, reduced range of C-arm angulations. Typical protocols for 3D C-arm CT acquisitions require projection images within an angular range of approximately 180 to 200 degrees. A Tomosynthesis image is generated using a range of only, e.g., 90 degrees, without a change in the sampling rate. The missing information introduces imaging artifacts in the reconstructed

image (see Fig. 4.16), but harmful radiation exposure can be significantly reduced. Another clinical benefit is the fact that the Tomosynthesis protocol requires less time for the intra-operative image acquisition and only a partial C-arm sweep.

Given the knowledge about the angles used to reconstruct the Tomosynthesis image, we infer the regions in the volume that contain artifacts. In the probability map, those regions are likely to be noisy and thus, the confidence of the individual probabilities is low. Therefore, we reduce the influence of the points $\vec{p}$ that lie within those regions by adapting their weight $w_{\vec{p}}$ accordingly. In practice, we compared each point in the ground-truth annotations of medical experts in corresponding (same patient) C-arm CT and Tomosynthesis images. Mesh points exhibiting high variance receive low weights, while points, whose annotations are similar get higher weights assigned. We show that using our framework and the adapted prior weights, the alignment of Tomosynthesis images with pre-operative images is possible and that we achieve similar registration performance as with standard 3D C-arm CT scans.

The major part of our experiments is based on regular C-arm CT images, since currently we only have access to a very limited number of pairs of corresponding CT and Tomosynthesis volumes. Figure 4.17 depicts two pairs of fused volumes. The images on the left show the registered CT and the complete 200 degree C-arm CT images. On the right, the corresponding results using the Tomosynthesis volumes (90 degrees angular acquisition range), instead of the regular C-arm CT images are shown.

4.6. Discussion

We presented two hybrid algorithms in order to align pre-operative CT and intra-operative 3D C-arm CT data in order to provide guidance to minimally-invasive cardiac procedures.

The first method uses the trachea bifurcation model as the anchor anatomy in order to align the pre-operative and intra-operative image. Complex acquisition protocols of contrasted 3D C-arm CT can be avoided as our method relies on the trachea bifurcation model which is visible in both modalities without adding contrast. Fast and robust machine learning algorithms are employed to estimate the final model parameters which can handle noisy intra-operative data. A weighted mapping transform is learned from training data to minimize the estimation error of the anatomy of interest, the aortic valve.

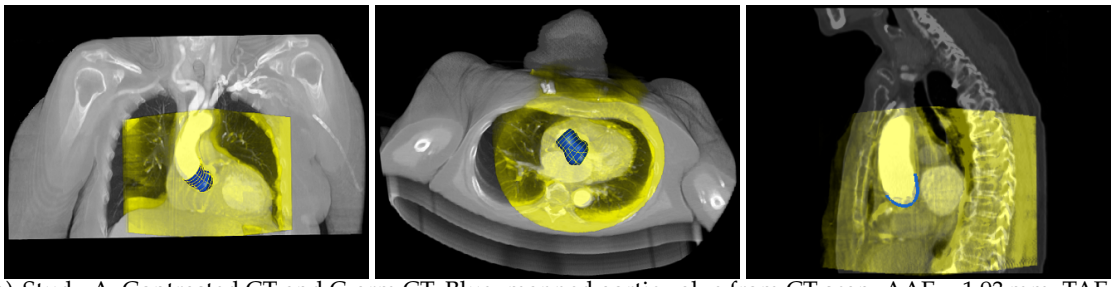
The second method represents a novel sparse matching approach to fuse the pre-operative anchor anatomy (in our case the pericardium) to the intra-operative setting. Data and model uncertainties are learned and exploited during matching. Quantitative and qualitative evaluation demonstrate a fast and accurate mapping of the anchor and target anatomy (in the case of the TAVI procedure the aortic valve model) to the intra-operative modality. In direct comparison with a state-of-the-art registration framework and a recently proposed quasi-global, knowledge-driven fusion approach, our method outperforms both in terms of robustness and accuracy regarding the targeted application. In addition the pericardium, used as an anchor anatomy, is in closer proximity as the trachea bifurcation

model (presented as the first method), we achieve more accurate results.

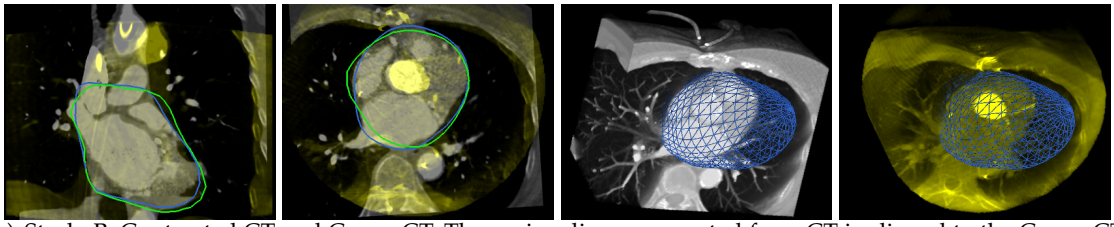
Furthermore, an inter-user variability study with ten users confirms that the accuracy of our method lies within the confidence interval of the expert group. While computation times of our method (1.6 s) and the quasi-global approach (2.5 s) are comparable, the ITK registration framework is significantly slower, typically consuming five to ten minutes per registration, similar to manual alignment done by experts.

The main limitation of our approach is the need for a large number of training datasets, such that the classifier is able to create reliable probability maps for C-arm CT images from a broad spectrum of potential scanners and acquisition protocols. Hence, a C-arm CT database with manual annotations of the anchor anatomy is necessary, which contains images acquired with and without contrast agent injected, images from different detectors (size and resolution) and various fields of view.

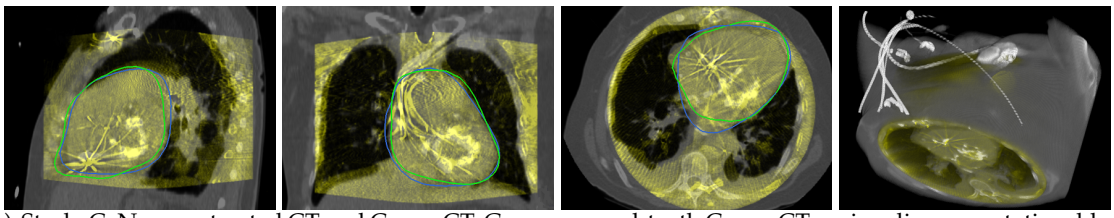
Comprehensive patient-specific models can be estimated from high-contrast CT and fused into the imaging environment of operating rooms to facilitate guidance in minimally-invasive cardiac surgery, while meeting interventionally necessary constraints such as low computation time, high accuracy and robustness against noisy data, partially visible models and imaging artifacts.



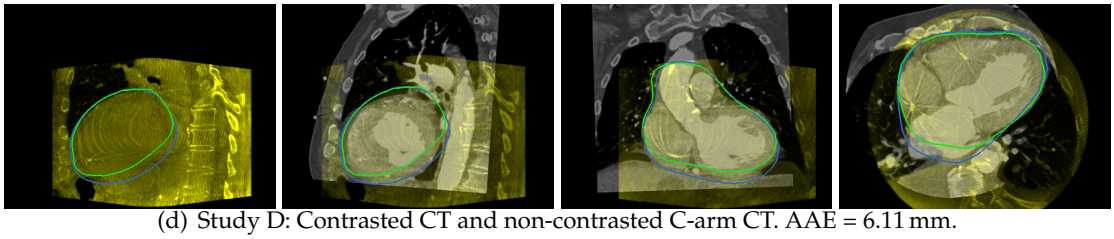
(a) Study A: Contrasted CT and C-arm CT. Blue: mapped aortic valve from CT scan. AAE = 1.93 mm, TAE = 1.24 mm.



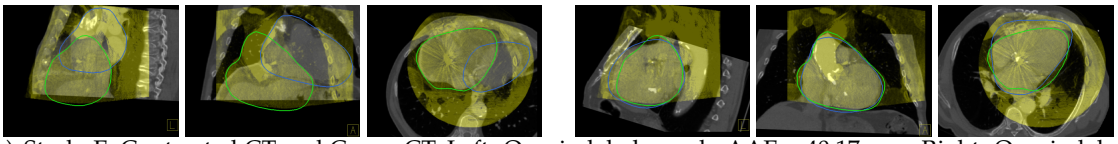
(b) Study B: Contrasted CT and C-arm CT. The pericardium segmented from CT is aligned to the C-arm CT. AAE = 4.07 mm.



(c) Study C: Non-contrasted CT and C-arm CT. Green: ground-truth C-arm CT pericardium annotation, blue: fused CT pericardium. AAE = 4.94 mm.



(d) Study D: Contrasted CT and non-contrasted C-arm CT. AAE = 6.11 mm.



(e) Study E: Contrasted CT and C-arm CT. Left: Quasi-global search, AAE = 40.17 mm. Right: Quasi-global search (masked), AAE = 5.24 mm.

Figure 4.15.: (a-d) Representative qualitative results from several automatically registered datasets. Yellow: intra-operative C-arm CT, gray: aligned high-quality pre-operative CT overlay, others: anatomical models automatically extracted (from CT scan) or annotated (based on C-arm CT), mapped into the joint coordinate system. The left image in the row (d) illustrates that especially in non-contrasted C-arm CT images, there are uncertainties involved in annotating the pericardium. It is not clear whether the fused or the annotated pericardium fits better. Non-ideal annotations usually increase the measured quantitative error. (e) Comparison method (Quasi-global Search) without and with pericardium masking. In this case, masking improves the result significantly.

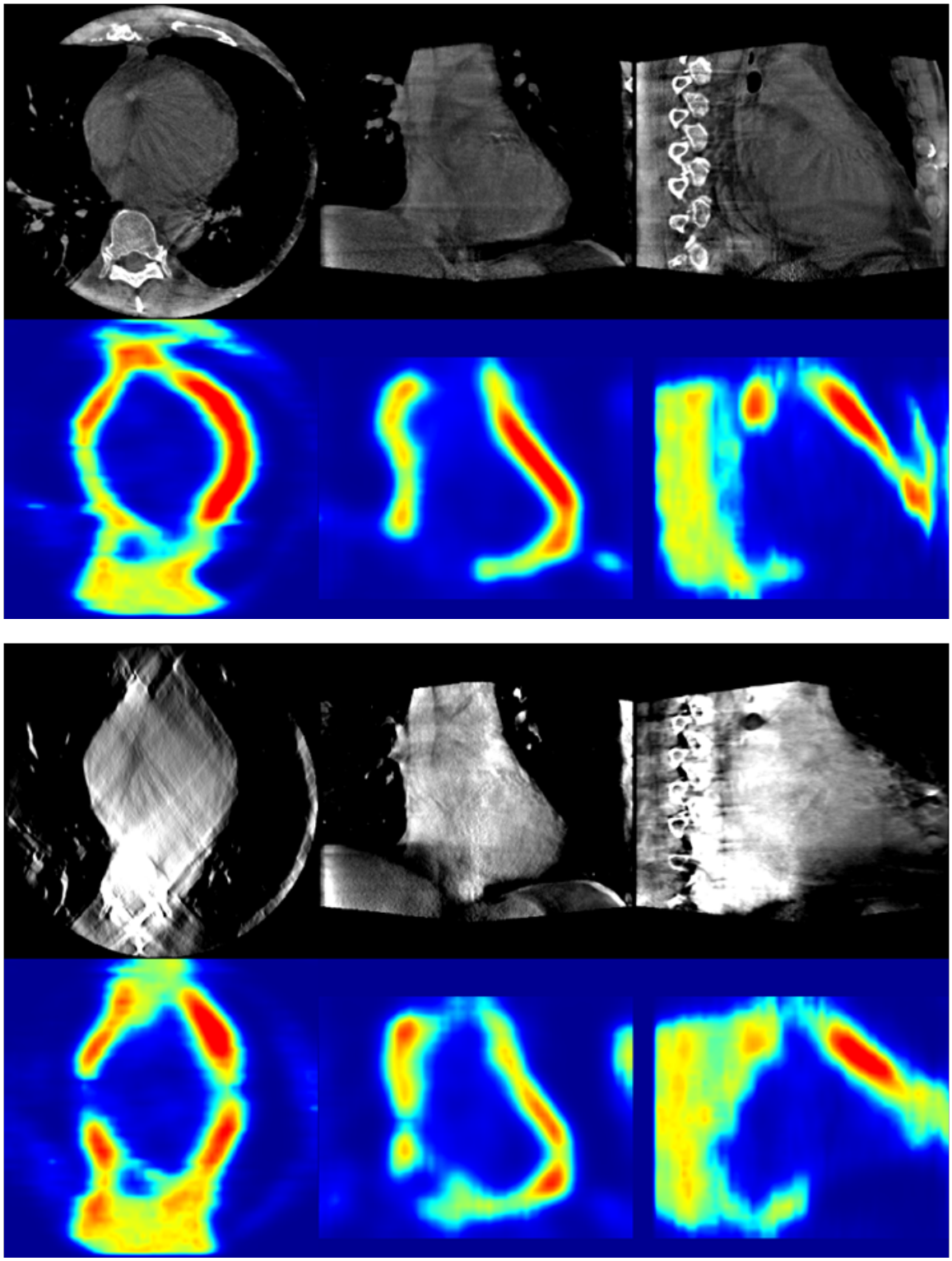


Figure 4.16.: Top: slices of volume reconstructed using regular 3D C-arm CT protocol (200 degrees) and corresponding probability map, bottom: slices of Tomosynthesis volume (90 degree C-arm sweep) and corresponding probability map

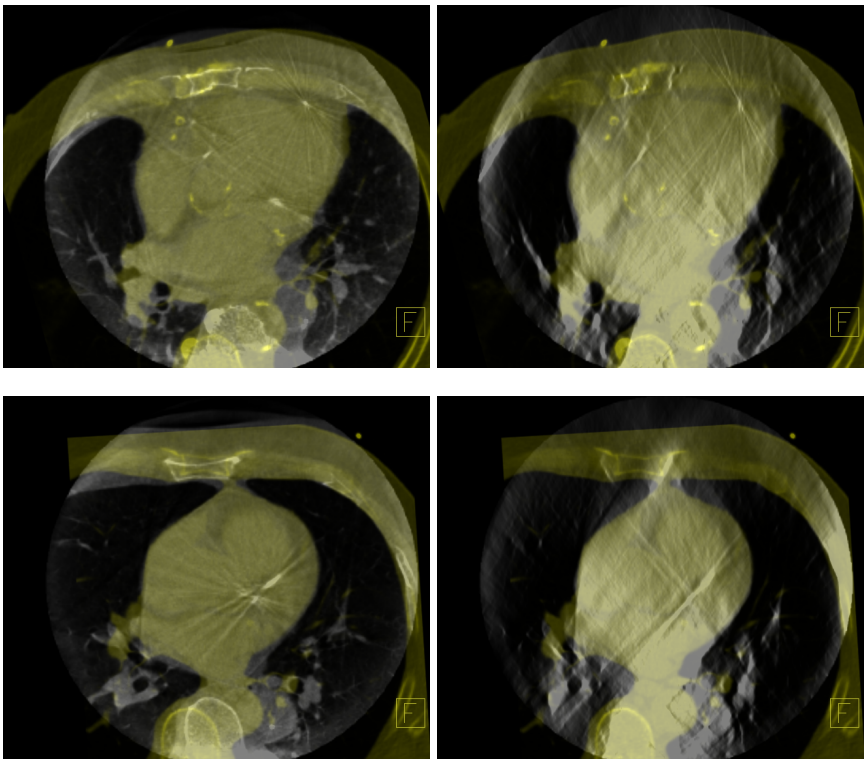


Figure 4.17.: Comparison of fusion of CT with C-arm CT (left), and CT with limited-angle Tomosynthesis images (right). The CT volume is depicted in yellow, the C-arm or Tomosynthesis image is depicted in grayscale.

5. Conclusion

5.1. Summary

The main focus of this thesis was to develop fast, precise and reproducible image analysis algorithms to assist the two main aspects of current minimally invasive valve procedures: the pre-operative planning and the intra-operative guidance.

Valvular Heart Disease (VHD) is the most prevalent subgroup of CVD affecting 2.5% of the global population and requiring yearly over 100,000 surgeries in the United States alone and is a representative instance for the growing public health problem provoked by cardiovascular diseases. Heart valve operations are the most expensive and the riskiest cardiac procedures, with an average cost of $141,120 and 4.9% in-hospital death rate. In chapter 2.8 we presented the current and upcoming non-invasive imaging modalities and their invaluable impact on current clinical practice for diagnosis, treatment planning and interventional guidance. We presented the limitation of the manual evaluation of these images whereby the complexity and non-reproducibility of the results is commonly observed. We advocated for the usage of advanced image analysis algorithms during the clinical workflow to circumvent these disadvantage and facilitate fast, precise and reproducible extraction of key measurements from image data in order to aid quantification, treatment planning and interventional guidance.

In chapter 3 we proposed a novel physiological model of all heart valves to precisely capture their morphological and dynamical properties. The model consists of three hierarchy layers (first a global motion model, second the anatomical landmark model and finally the full surface model). The hierarchical parametrization of our model allows for the usage of efficient parameter estimation techniques. The algorithm used to estimate these parameters is based on robust and fast machine learning algorithms in combination with a constrained multi-linear shape space and enables patient specific model estimation within three minutes and an accuracy of 1.24 mm from a multi-phase CT data set. Further the automatic estimation error of the valve models are slightly above (1.13 mm for the aortic valve and 0.78 mm for the mitral valve) the inter-user variability. Several extensions were proposed. In section 3.6 a new volumetric representation of the aortic valve was presented and the different tissues within the valve were classified (calcium, leaflet tissue and blood pool) using a multi-label boosting classifier and an ensemble of geometric and intensity based features. We further showed that based on the estimated volumetric model several clinical applications can be supported, such as intervention planning for TAVI. In section 3.7 we demonstrated that by using our estimated aortic valve model and a standard simulation framework we can efficiently simulate device deployment for the

TAVI procedure. The technology described can potentially advance the management of patients affected by valvular heart disease by reducing clinical diagnosis costs (due to the automation and reproducibility of results which could be obtained with our robust model estimation techniques) and reduce risk of complications during minimally invasive procedures as advanced assessment and intervention planning can be performed before the procedure.

Finally, in chapter 4 we presented a robust model based algorithm to align multi-modal data. First we introduced a method where we first estimated geometric models using fast and robust machine learning techniques from two modalities independently. Afterward the alignment is done using a weighted mapping transform which is learned from training in order to achieve optimal registration results. Second we introduce an improved version where the complete model estimation is not necessary in the intra-operative modality. Using a novel sparse matching algorithm and a trained likelihood classifier to highlight the anatomy of interest in the intra-operative image we can bring both modalities into correspondence. The algorithm is ideally suited to the clinical application of guidance for minimally invasive procedures.

5.2. Future Work

Physiological Modeling: The first option would be to build a generic representation consisting of all cardiac anatomies. Thus the current model parametrization must be extended to the heart chambers (left ventricle, left atrium, right ventricle and right atrium). Other options would be to develop a better model parametrization of the current valve models. More details of the valvular anatomy could be captured in this representations such as the sub-valvular apparatus of the mitral valve, consisting of the papillary muscles and the chordae tendae. In addition a volumetric representation of all the valves would be helpful. Especially in the context of biomechanical simulation of valve repair, implant deployments etc. it would provide more accurate boundary conditions than generic thickness values.

Non-linear Model-Based Fusion: The proposed Model-Based Fusion approach presented in chapter 4 is estimating a rigid transformation in order to align two modalities. This method could be extended to allow for the estimation of non-linear deformations. Therefore an updated transformation model must be used in combination with an efficient optimization algorithm. Hereby the optimization should not be done using the image intensities but rather using the more robust probability map which is generated by training a classifier to enhance important structural information used during the matching procedure and suppress anatomies which are not of interest.

3D-2D Model-Based Fusion: Using the Model-Based Fusion approach we could align pre-operative modalities with the intra-operative 2D Fluoroscopic images. Thus a patient specific anatomical model could be extracted from pre-operative data and integrated in the intra-operative setting. It can be seen as a continuation of the method proposed in section 4 to align a pre-operative image with mono and bi-plane fluoroscopic images. Thus

the complex protocol to acquire the 3D C-arm CT used for registration can be avoided. In addition to the initial alignment there is a continuous need to adjust the target anatomy model used during guidance due to the breathing and cardiac motion during the intervention. An algorithm must be developed which would compensate the current anatomical model for cardiac and breathing motion in real-time.

Appendix

A. List of Authored Publications and Patents

i. D. NEUMANN, S. GRBIC; R. IONASEC, M. JOHN, N. NAVAB, J. HORNEGGER, *Robust Model-based 3D/3D Fusion using Sparse Matching for Minimally Invasive Surgery*, IEEE Transactions on Medical Imaging - TMI, submitted in August 2013.

ii. S. GRBIC, T. MANSI, R. IONASEC, I. VOIGT, H. HOULE, M. JOHN, M. SCHOEBINGER, N. NAVAB, D. COMANICIU, *Image-Based Computational Models for TAVI Planning: From CT Images to Implant Deployment*, Medical Image Computing and Computer Assisted Intervention - MICCAI 2013, Nagoya, Japan, September 2013.

iii. D. NEUMANN, S. GRBIC, M. JOHN, N. NAVAB, J. HORNEGGER, R. IONASEC, *Robust Model-based 3D/3D Fusion using Sparse Matching for Minimally Invasive Surgery*, Medical Image Computing and Computer Assisted Intervention - MICCAI 2013, Nagoya, Japan, September 2013.

iv. S. GRBIC, R. IONASEC, T. MANSI, B. GEORGESCU, F. VEGA-HIGUERA, N. NAVAB, D. COMANICIU, *Advanced Intervention planning for Transcatheter Aortic Valve Implantations (TAVI) from CT using volumetric models*, IEEE International Symposium on Biomedical Imaging - ISBI 2013, San Francisco, USA, April 2013.

v. S. GRBIC, R. IONASEC, D. VITANOVSKI, I. VOIGT, Y. WANG, B. GEORGESCU, N. NAVAB, D. COMANICIU, *Complete Valvular Heart Apparatus from 4D Cardiac CT*, Medical Image Analysis - MedIA, March 2012.

vi. S. GRBIC, C. GESELL, R. IONASEC, M. JOHN, J. BOESE, N. NAVAB, D. COMANICIU, *Model-Based Fusion of CT and non-contrasted 3D C-arm CT: Application to Transcatheter Valve Therapies*, IEEE International Symposium on Biomedical Imaging - ISBI 2012, Barcelona, Spain, May 2012.

vii. S. GRBIC, R. IONASEC, Y. WANG, T. MANSI, B. GEORGESCU, M. JOHN, J. BOESE, Y. ZHENG, N. NAVAB, D. COMANICIU, *Model-Based Fusion of Multi-Modal Volumetric Images: Application to Transcatheter Valve Procedures*, Medical Image Computing and Computer Assisted Intervention - MICCAI 2011, Toronto, Canada, September 2011.

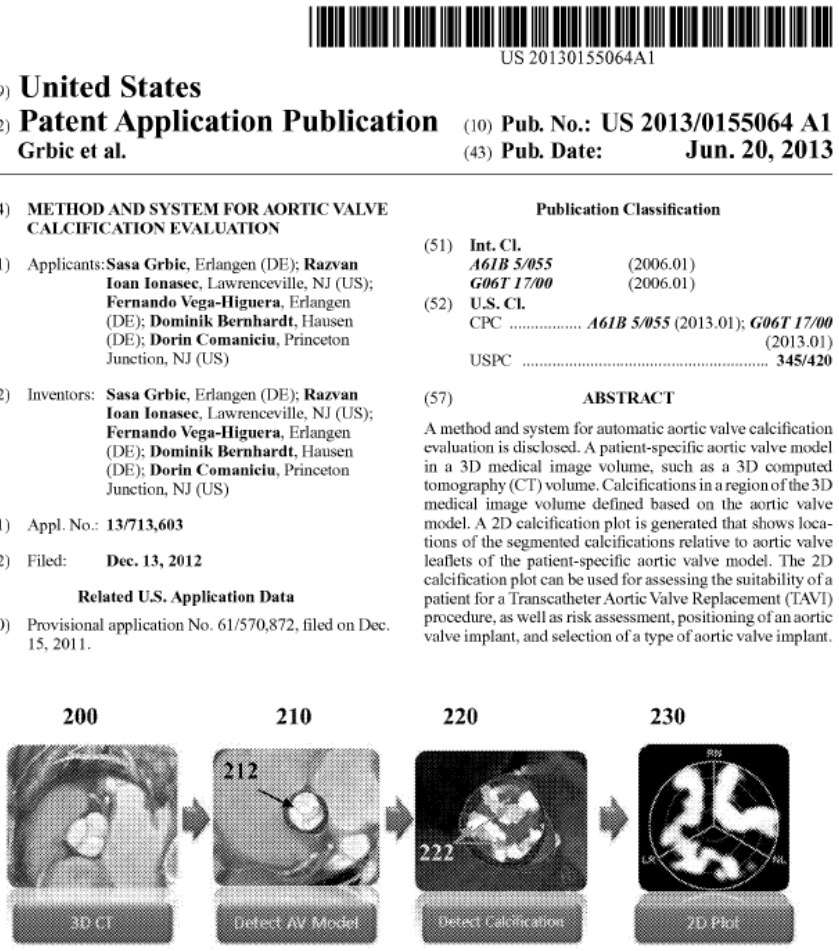
viii. IONASEC, I. VOIGT, V. MIHALEF, S. GRBIC, D. VITANOVSKI, Y. WANG, Y. ZHENG, J. HORNEGGER, N. NAVAB, B. GEORGESCU, D. COMANICIU, *Patient-specific Modeling of the Heart: Applications to Cardiovascular Disease Management*, Medical Image Computing and Computer Assisted Intervention: Workshop on Statistical Atlases and

Computational Models of the Heart: Mapping Structure and Function - MICCAI 2010, Beijing, China, September 2010.

ix. S. GRBIC, R. IONASEC, D. VITANOVSKI, I. VOIGT, B. GEORGESCU, , N. NAVAB, D. COMANICIU, *Complete Valvular Heart Apparatus Model from 4D Cardiac CT*, Medical Image Computing and Computer Assisted Intervention - MICCAI 2010, Beijing, China, September 2010.

x. S. GRBIC, R. IONASEC, Y. ZHENG, D. ZAEUNER, B. GEORGESCU, D. COMANICIU, *Aortic Valve and Ascending Aortic Root Modeling from 3D and 3D+t CT*, SPIE Medical Imaging, 2010, San Diego, USA, February 2010.

xi. S. GRBIC, M. URSCHLER, H. BISCHOF, *Optical Flow based Deformable Volume Registration using a novel Second-Order Regularization Prior*, SPIE Medical Imaging, 2010, San Diego, USA, February 2010.



US 20130155064A1

(19) **United States**

(12) **Patent Application Publication**
Grbic et al.

(10) **Pub. No.: US 2013/0155064 A1**
(43) **Pub. Date: Jun. 20, 2013**

(54) **METHOD AND SYSTEM FOR AORTIC VALVE CALCIFICATION EVALUATION**

(71) Applicants: **Sasa Grbic**, Erlangen (DE); **Razvan Ioan Ionasec**, Lawrenceville, NJ (US); **Fernando Vega-Higuera**, Erlangen (DE); **Dominik Bernhardt**, Hausen (DE); **Dorin Comaniciu**, Princeton Junction, NJ (US)

(72) Inventors: **Sasa Grbic**, Erlangen (DE); **Razvan Ioan Ionasec**, Lawrenceville, NJ (US); **Fernando Vega-Higuera**, Erlangen (DE); **Dominik Bernhardt**, Hausen (DE); **Dorin Comaniciu**, Princeton Junction, NJ (US)

(21) Appl. No.: **13/713,603**

(22) Filed: **Dec. 13, 2012**

Related U.S. Application Data

(60) Provisional application No. 61/570,872, filed on Dec. 15, 2011.

Publication Classification

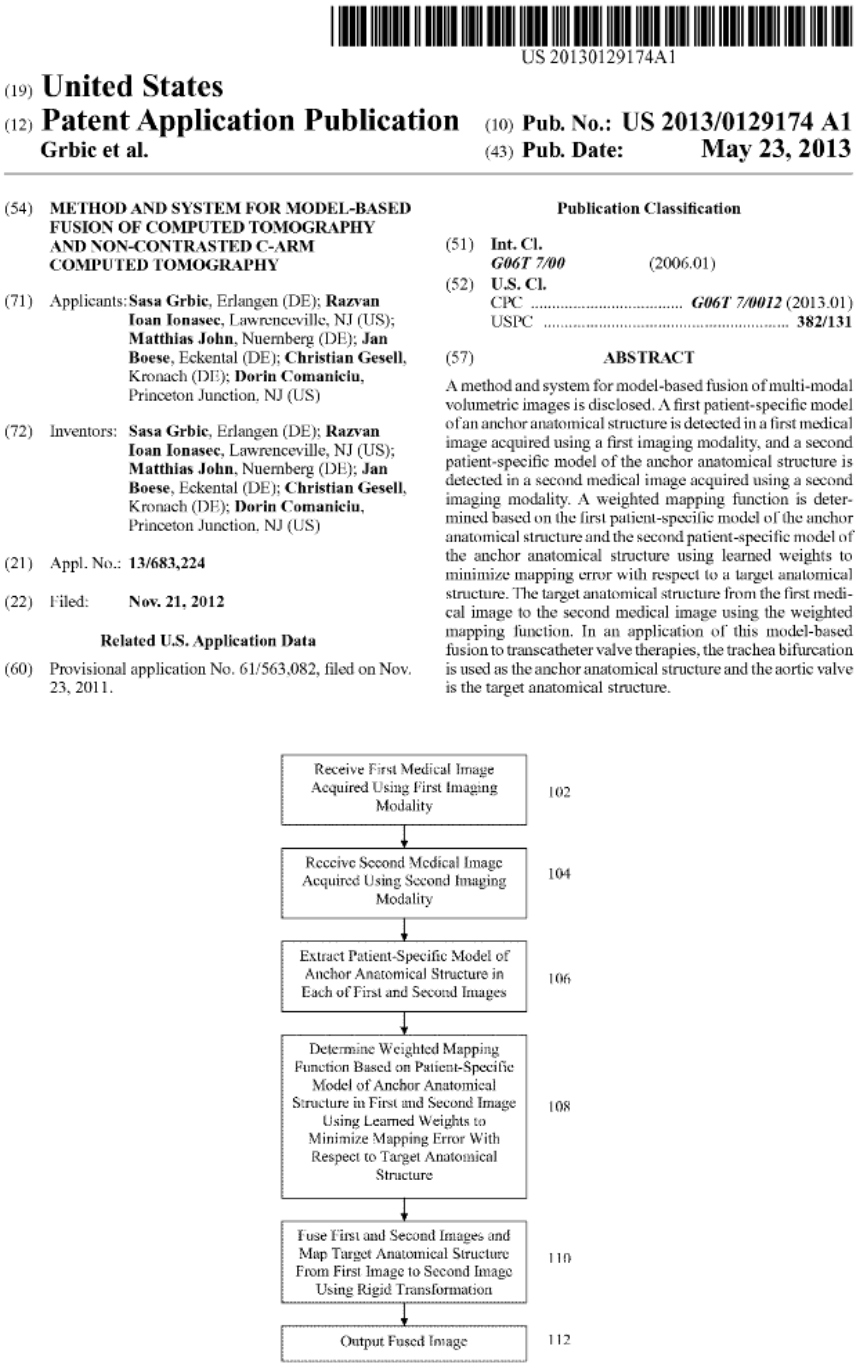
(51) **Int. Cl.**
A61B 5/055 (2006.01)
G06T 17/00 (2006.01)

(52) **U.S. Cl.**
CPC ***A61B 5/055*** (2013.01); ***G06T 17/00*** (2013.01)
USPC **345/420**

(57) **ABSTRACT**

A method and system for automatic aortic valve calcification evaluation is disclosed. A patient-specific aortic valve model in a 3D medical image volume, such as a 3D computed tomography (CT) volume. Calcifications in a region of the 3D medical image volume defined based on the aortic valve model. A 2D calcification plot is generated that shows locations of the segmented calcifications relative to aortic valve leaflets of the patient-specific aortic valve model. The 2D calcification plot can be used for assessing the suitability of a patient for a Transcatheter Aortic Valve Replacement (TAVI) procedure, as well as risk assessment, positioning of an aortic valve implant, and selection of a type of aortic valve implant.

Figure A.1.: Patent US 2013/0155064, http://www.freepatentsonline.com/20130155064.pdf



US 20130129174A1

(19) **United States**

(12) **Patent Application Publication**

Grbic et al.

(10) **Pub. No.: US 2013/0129174 A1**

(43) **Pub. Date: May 23, 2013**

(54) **METHOD AND SYSTEM FOR MODEL-BASED FUSION OF COMPUTED TOMOGRAPHY AND NON-CONTRASTED C-ARM COMPUTED TOMOGRAPHY**

(71) Applicants: **Sasa Grbic**, Erlangen (DE); **Razvan Ioan Ionasec**, Lawrenceville, NJ (US); **Matthias John**, Nuernberg (DE); **Jan Boese**, Eckental (DE); **Christian Gesell**, Kronach (DE); **Dorin Comaniciu**, Princeton Junction, NJ (US)

(72) Inventors: **Sasa Grbic**, Erlangen (DE); **Razvan Ioan Ionasec**, Lawrenceville, NJ (US); **Matthias John**, Nuernberg (DE); **Jan Boese**, Eckental (DE); **Christian Gesell**, Kronach (DE); **Dorin Comaniciu**, Princeton Junction, NJ (US)

(21) Appl. No.: **13/683,224**

(22) Filed: **Nov. 21, 2012**

Related U.S. Application Data

(60) Provisional application No. 61/563,082, filed on Nov. 23, 2011.

Publication Classification

(51) **Int. Cl.**
G06T 7/00 (2006.01)

(52) **U.S. Cl.**
CPC ***G06T 7/0012*** (2013.01)
USPC **382/131**

(57) **ABSTRACT**

A method and system for model-based fusion of multi-modal volumetric images is disclosed. A first patient-specific model of an anchor anatomical structure is detected in a first medical image acquired using a first imaging modality, and a second patient-specific model of the anchor anatomical structure is detected in a second medical image acquired using a second imaging modality. A weighted mapping function is determined based on the first patient-specific model of the anchor anatomical structure and the second patient-specific model of the anchor anatomical structure using learned weights to minimize mapping error with respect to a target anatomical structure. The target anatomical structure from the first medical image to the second medical image using the weighted mapping function. In an application of this model-based fusion to transcatheter valve therapies, the trachea bifurcation is used as the anchor anatomical structure and the aortic valve is the target anatomical structure.

Figure A.2.: Patent US 2013/0129174, http://www.freepatentsonline.com/20130129174.pdf

US 20130129173A1

(19) **United States**

(12) **Patent Application Publication**

Grbic et al.

(10) **Pub. No.: US 2013/0129173 A1**

(43) **Pub. Date: May 23, 2013**

(54) **METHOD AND SYSTEM FOR INTERVENTION PLANNING FOR TRANSCATHETER AORTIC VALVE IMPLANTATION FROM 3D COMPUTED TOMOGRAPHY DATA**

(71) Applicants: **Sasa Grbic**, Erlangen (DE); **Razvan Ioan Ionasec**, Lawrenceville, NJ (US); **Fernando Vega-Higuera**, Erlangen (DE); **Dominik Bernhardt**, Hausen (DE); **Bogdan Georgescu**, Plainsboro, NJ (US); **Dorin Comaniciu**, Princeton Junction, NJ (US)

(72) Inventors: **Sasa Grbic**, Erlangen (DE); **Razvan Ioan Ionasec**, Lawrenceville, NJ (US); **Fernando Vega-Higuera**, Erlangen (DE); **Dominik Bernhardt**, Hausen (DE); **Bogdan Georgescu**, Plainsboro, NJ (US); **Dorin Comaniciu**, Princeton Junction, NJ (US)

(21) Appl. No.: **13/682,820**

(22) Filed: **Nov. 21, 2012**

Related U.S. Application Data

(60) Provisional application No. 61/563,090, filed on Nov. 23, 2011.

Publication Classification

(51) **Int. Cl.**
G06T 7/00 (2006.01)

(52) **U.S. Cl.**
CPC ***G06T 7/0012*** (2013.01)
USPC **382/131**

(57) **ABSTRACT**

A method and system for automated intervention planning for transcatheter aortic valve implantations using computed tomography (CT) data is disclosed. A patient-specific aortic valve model is detected in a CT volume of a patient. The patient-specific aortic valve model is detected by detecting a global location of the patient-specific aortic valve model in the CT volume, detecting aortic valve landmarks based on the detected global location, and fitting an aortic root surface model. Angulation parameters of a C-arm imaging device for acquiring intra-operative fluoroscopic images and anatomical measurements of the aortic valve are automatically determined based on the patient-specific aortic valve model.

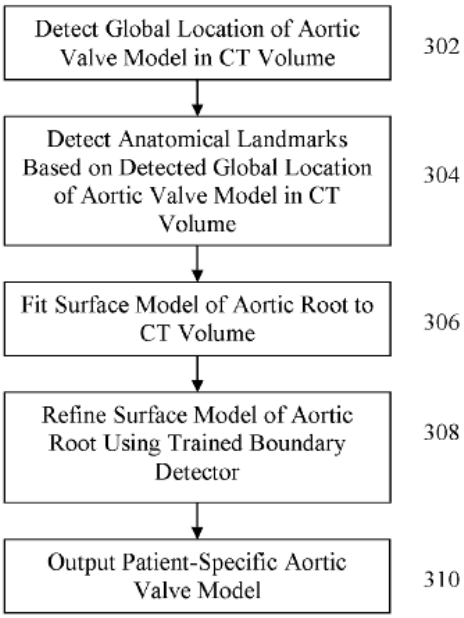


Figure A.3.: Patent US 2013/0129173, http://www.freepatentsonline.com/20130129173.pdf

US 20120230568A1

(19) **United States**

(12) **Patent Application Publication**
Grbic et al.

(10) **Pub. No.: US 2012/0230568 A1**

(43) **Pub. Date: Sep. 13, 2012**

(54) **METHOD AND SYSTEM FOR MODEL-BASED FUSION OF MULTI-MODAL VOLUMETRIC IMAGES**

(75) Inventors: **Sasa Grbic**, Erlangen (DE); **Razvan Ioan Ionasec**, Lawrenceville, NJ (US); **Yang Wang**, Princeton, NJ (US); **Bogdan Georgescu**, Plainsboro, NJ (US); **Tommaso Mansi**, Westfield, NJ (US); **Dorin Comaniciu**, Princeton Junction, NJ (US); **Yefeng Zheng**, Dayton, NJ (US); **Shaohua Kevin Zhou**, Plainsboro, NJ (US); **Matthias John**, Nurnberg (DE); **Jan Boese**, Eckental (DE)

(73) Assignees: **Siemens Aktiengesellschaft**, Munich (DE); **Siemens Corporation**, Iselin, NJ (US)

(21) Appl. No.: **13/412,869**

(22) Filed: **Mar. 6, 2012**

Related U.S. Application Data

(60) Provisional application No. 61/451,006, filed on Mar. 9, 2011.

Publication Classification

(51) **Int. Cl.**
G06K 9/36 (2006.01)

(52) **U.S. Cl.** **382/131**; 382/128

(57) **ABSTRACT**

A method and system for fusion of multi-modal volumetric images is disclosed. A first image acquired using a first imaging modality is received. A second image acquired using a second imaging modality is received. A model and of a target anatomical structure and a transformation are jointly estimated from the first and second images. The model represents a model of the target anatomical structure in the first image and the transformation projects a model of the target anatomical structure in the second image to the model in the first image. The first and second images can be fused based on estimated transformation.

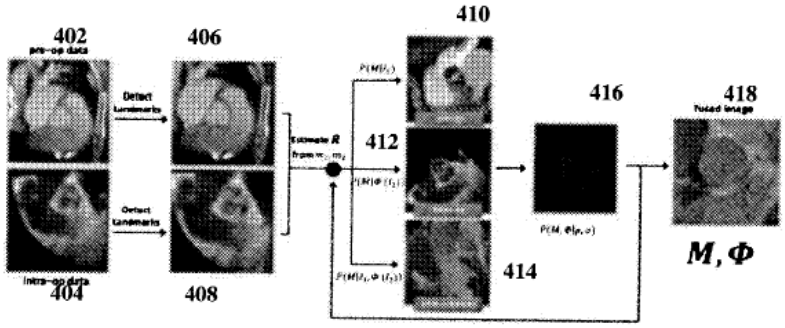


Figure A.4.: Patent US 2012/0230568, http://www.freepatentsonline.com/20120230568.pdf

US 20120022843A1

(19) **United States**

(12) **Patent Application Publication**

Ionasec et al.

(10) **Pub. No.: US 2012/0022843 A1**

(43) **Pub. Date: Jan. 26, 2012**

(54) **METHOD AND SYSTEM FOR COMPREHENSIVE PATIENT-SPECIFIC MODELING OF THE HEART**

(76) Inventors: **Razvan Ioan Ionasec**, Lawrenceville, NJ (US); **Ingmar Voigt**, Erlangen (DE); **Viorel Mihalef**, Keasbey, NJ (US); **Sasa Grbic**, Erlangen (DE); **Dime Vitanovski**, Erlangen (DE); **Yang Wang**, Plainsboro, NJ (US); **Yefeng Zheng**, Dayton, NJ (US); **Bogdan Georgescu**, Plainsboro, NJ (US); **Dorin Comaniciu**, Princeton Junction, NJ (US); **Puneet Sharma**, Rahway, NJ (US); **Tommaso Mansi**, Westfield, NJ (US)

(21) Appl. No.: **13/091,076**

(22) Filed: **Apr. 20, 2011**

Related U.S. Application Data

(60) Provisional application No. 61/366,294, filed on Jul. 21, 2010, provisional application No. 61/383,942, filed on Sep. 17, 2010, provisional application No. 61/409,633, filed on Nov. 3, 2010.

Publication Classification

(51) **Int. Cl.**
G06G 7/60 (2006.01)
G06G 7/57 (2006.01)

(52) **U.S. Cl.** **703/9**; 703/11

(57) **ABSTRACT**

A method and system for patient-specific modeling of the whole heart anatomy, dynamics, hemodynamics, and fluid structure interaction from 4D medical image data is disclosed. The anatomy and dynamics of the heart are determined by estimating patient-specific parameters of a physiological model of the heart from the 4D medical image data for a patient. The patient-specific anatomy and dynamics are used as input to a 3D Navier-Stokes solver that derives realistic hemodynamics, constrained by the local anatomy, along the entire heart cycle. Fluid structure interactions are determined iteratively over the heart cycle by simulating the blood flow at a given time step and calculating the deformation of the heart structure based on the simulated blood flow, such that the deformation of the heart structure is used in the simulation of the blood flow at the next time step. The comprehensive patient-specific model of the heart representing anatomy, dynamics, hemodynamics, and fluid structure interaction can be used for non-invasive assessment and diagnosis of the heart, as well as virtual therapy planning and cardiovascular disease management. Parameters of the comprehensive patient-specific model are changed or perturbed to simulate various conditions or treatment options, and then the patient specific model is recalculated to predict the effect of the conditions or treatment options.

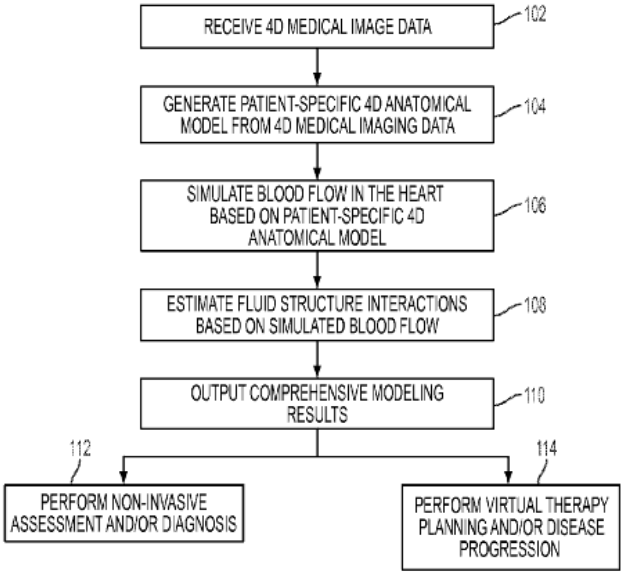


Figure A.5.: Patent US 2012/0022843, http://www.freepatentsonline.com/20120022843.pdf

US 20130035596A1

(19) **United States**

(12) **Patent Application Publication**

Ionasec et al.

(10) **Pub. No.: US 2013/0035596 A1**

(43) **Pub. Date: Feb. 7, 2013**

(54) **MODEL-BASED POSITIONING FOR INTRACARDIAC ECHOCARDIOGRAPHY VOLUME STITCHING**

(75) Inventors: **Razvan Ioan Ionasec**, Lawrenceville, NJ (US); **Sasa Grbic**, Erlangen (DE); **Estelle Camus**, Mountain View, CA (US); **Dorin Comaniciu**, Princeton Junction, NJ (US)

(73) Assignee: **Siemens Corporation**, Iselin, NJ (US)

(21) Appl. No.: **13/546,066**

(22) Filed: **Jul. 11, 2012**

Related U.S. Application Data

(60) Provisional application No. 61/507,703, filed on Jul. 14, 2011.

Publication Classification

(51) **Int. Cl.**
A61B 8/12 (2006.01)

(52) **U.S. Cl.** **600/450**

(57) **ABSTRACT**

Different intracardiac echocardiography volumes are stitched together. Different volumes of a patient are scanned with ICE. To stitch the volumes together, creating a larger volume, the volumes are spatially aligned. The alignment is based on feature, surface, or both feature and surface matching of the ICE volumes with a preoperative model of the same patient. The matching with the model indicates a relative position of the ICE volumes with each other. Using machine-trained classifiers may speed performance, allowing for real-time assembling of a volume from ICE data as the catheter is moved within the patient.

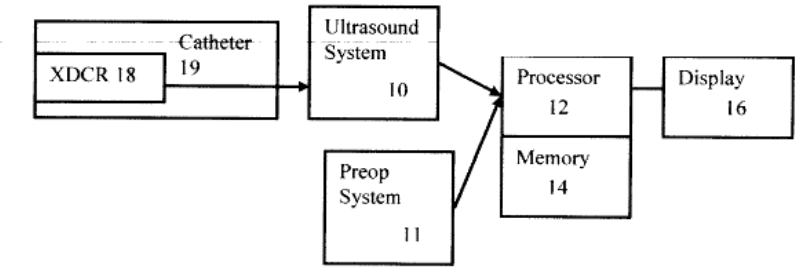


Figure A.6.: Patent US 2013/0035596, http://www.freepatentsonline.com/20130035596.pdf

Glossary

AV Aortic Valve. 2, 3, 63, 64, 100, 108, 111, 142

CT Computed Tomography. 2–4, 15, 19–21, 23, 26–31, 33–36, 53, 58, 59, 61, 62, 68, 70–76, 82–85, 91, 92, 94, 96–102, 108, 110, 112, 113, 115, 117, 118, 123, 139, 140, 142–144

CVD Cardiovascular Disease. 1

ECG Electrocardiography. 7, 20, 29, 58

GPA Generalized Procrustes Analysis. 51, 52, 55

ICA Independent Component Analysis. 61

ITK Insight Segmentation and Registration Toolkit. 112

IVUS Intravascular ultrasound. 24

MI Mutual Information. 94

MICA Multi-linear Independent Component Analysis. 55

MPCA Multi-linear Principal Component Analysis. 55

MRI Magnetic Resonance Imaging. 22, 23, 26–28, 85, 91, 92, 94, 140

MSL Marginal Space Learning. 50, 51, 101

MV Mitral Valve. 2, 3, 27, 29, 30, 84

NCC Normalized Cross Correlation. 93

OCT Optical Coherence Tomography. 24, 25, 140

PBT Probabilistic Boosting Tree. 40, 43, 77, 80, 97, 101, 103

PCA Principal Component Analysis. 52, 56, 61, 81

PV Pulmonary Valve. 2, 3, 63, 84, 142

SAD Sum of absolute differences. 92, 93

SSD Sum of squared differences. 92

TAVI Transcatheter Aortic Valve Implantation. 3, 31, 63, 83

TEE Transesophageal Echocardiography. 2, 18, 22, 23, 26–28, 30, 76, 82–85, 140, 143

TTE Transthoracic Echocardiography. 22, 23, 140

TV Tricuspid Valve. 2, 3, 84

VHD Valvular Heart Disease. 1, 123

List of Figures

Bibliography

[1] Ankush Aggarwal, Vanessa Aguilar, Chung-Hao Lee, Giovanni Ferrari, Joseph Gorman, Rober Gorman, and Michael Sacks. Patient-specific modeling of heart valves: From image to simulation. In Sebastien Ourselin, Daniel Rueckert, and Nicolas Smith, editors, *Functional Imaging and Modeling of the Heart*, volume 7945 of *Lecture Notes in Computer Science*, pages 141–149. Springer Berlin Heidelberg, 2013.

[2] Simon L Altmann. *Rotations, quaternions, and double groups*. Dover Publications, 2005.

[3] A. Andreopoulos and J. K. Tsotsos. A novel algorithm for fitting 3-D active appearance models: Application to cardiac MRI segmentation. In *Proc. Scandinavian Conf. Image Analysis*, pages 729–739, 2005.

[4] Elizabeth Arias. United states life tables, 2003. *National vital statistics reports: from the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System*, 54(14):1, 2006.

[5] Michel A Audette, Frank P Ferrie, and Terry M Peters. An algorithmic overview of surface registration techniques for medical imaging. *Medical Image Analysis*, 4(3):201–217, 2000.

[6] Yonatan Aumann and Yuval Rabani. An o (log k) approximate min-cut max-flow theorem and approximation algorithm. *SIAM Journal on Computing*, 27(1):291–301, 1998.

[7] WG Austen, JE Edwards, RL Frye, GG Gensini, VL Gott, LS Griffith, DC McGoon, ML Murphy, and BB Roe. A reporting system on patients evaluated for coronary artery disease. report of the ad hoc committee for grading of coronary artery disease, council on cardiovascular surgery, american heart association. *Circulation*, 51(4):5–40, 1975.

[8] Brian B Avants, Charles L Epstein, Murray Grossman, and James C Gee. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical image analysis*, 12(1):26–41, 2008.

[9] Z. Bao, L. Zhukov, I. Guskov, J. Wood, and D. Breen. Dynamic deformable models for 3D MRI heart segmentation. In *SPIE Medical Imaging*, pages 398–405, 2002.

[10] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. In *Proc. European Conf. Computer Vision*, pages 404–417, 2006.

[11] Jacob Beutel, Harold L Kundel, and Richard L Van Metter. Handbook of medical imaging, volume 1: Physics and psychophysics. 2000.

[12] Linda K Bickerstaff, Peter C Pairolero, Larry H Hollier, L Joseph Melton, Hubert J Van Peenen, and Kenneth J Cherry. Thoracic aortic aneurysms: a population-based study. *Atherosclerosis*, 15:29, 1982.

[13] Rémi Blanc, Mauricio Reyes, Christof Seiler, and Gábor Székely. Conditional variability of statistical shape models based on surrogate variables. In Guang-Zhong Yang, David Hawkes, Daniel Rueckert, Alison Noble, and Chris Taylor, editors, *Medical Image Computing and Computer Assisted Intervention*, volume 5762 of *Lecture Notes in Computer Science*, pages 84–91. Springer Berlin / Heidelberg, 2009.

[14] R. O. Bonow, B. A. Carabello, K. Chatterjee, A. C. Jr. de Leon, D. P. Faxon, M. D. Freed, W. H. Gaasch, B. W. Lytle, R. A. Nishimura, P. T. Gara, R. A. Rourke, C. M. Otto, P. M. Shah, and J. S. Shanewise. Acc/aha 2006 guidelines for the management of patients with valvular heart disease: a report of the american college of cardiology/american heart association task force on practice guidelines (writing committee to develop guidelines for the management of patients with valvular heart disease). *Circulation*, 114(5):84–231, 2006.

[15] Robert O Bonow, Blase A Carabello, Kanu Chatterjee, Antonio C de Leon, David P Faxon, Michael D Freed, William H Gaasch, Bruce W Lytle, Rick A Nishimura, Patrick T O'Gara, et al. 2008 focused update incorporated into the acc/aha 2006 guidelines for the management of patients with valvular heart diseasea report of the american college of cardiology/american heart association task force on practice guidelines (writing committee to revise the 1998 guidelines for the management of patients with valvular heart disease) endorsed by the society of cardiovascular anesthesiologists, society for cardiovascular angiography and interventions, and society of thoracic surgeons. *Journal of the American College of Cardiology*, 52(13):e1–e142, 2008.

[16] Guillaume Bouchard, Bill Triggs, et al. The tradeoff between generative and discriminative classifiers. In *IASC International Symposium on Computational Statistics (COMPSTAT)*, pages 721–728, 2004.

[17] Yuri Boykov and Vladimir Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(9):1124–1137, 2004.

[18] Yuri Boykov, Olga Veksler, and Ramin Zabih. Markov random fields with efficient approximations. In *Computer vision and pattern recognition, 1998. Proceedings. 1998 IEEE computer society conference on*, pages 648–655. IEEE, 1998.

[19] Yuri Boykov, Olga Veksler, and Ramin Zabih. Fast approximate energy minimization via graph cuts. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(11):1222–1239, 2001.

[20] Yuri Y Boykov and M-P Jolly. Interactive graph cuts for optimal boundary & region segmentation of objects in nd images. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 1, pages 105–112. IEEE, 2001.

[21] Chaim Broit. *Optimal registration of deformed images*. PhD thesis, University of Pennsylvania, 1981.

[22] Alexander M Bronstein, Michael, Computer Science, Technion Israel, Fabrice Michel, Nikos Paragios, Equipe Galen, and Inria Saclay Ile-de france. Data Fusion through Cross-modality Metric Learning using Similarity-Sensitive Hashing. *CVPR*, pages 3594 – 3601, 2010.

[23] P. Burlina, C. Sprouse, D. DeMenthon, A. Jorstad, R. Juang, F. Contijoch, T. Abraham, D. Yuh, and E. McVeigh. Patient-specific modeling and analysis of the mitral valve using 3d-tee. *Information Processing in Computer-Assisted Interventions*, pages 135–146, 2010.

[24] C Capelli, GM Bosi, E Cerri, J Nordmeyer, T Odenwald, P Bonhoeffer, F Migliavacca, AM Taylor, and S Schievano. Patient-specific simulations of transcatheter aortic valve stent implantation. *Medical and Biological Engineering and Computing*, pages 1–10, 2012.

[25] Blase A Carabello. Mitral valve regurgitation. *Current problems in cardiology*, 23(4):197–241, 1998.

[26] Alain Carpentier, David Adams, and Farzan Filsoufi. *Carpentier's Reconstructive Valve Surgery*. Saunders, 2010.

[27] Barbara Chapman, Gabriele Jost, and Ruud van der Pas. *Using OpenMP: Portable Shared Memory Parallel Programming (Scientific and Engineering Computation)*. The MIT Press, Cambridge, MA (US), first edition, 2007.

[28] A. Cheng, P. Dagum, and DC. Miller. Aortic root dynamics and surgery: from craft to science. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 362(1484):1407–1419, 2007.

[29] W Randolph Chitwood Jr, Joseph R Elbeery, and Jon F Moran. Minimally invasive mitral valve repair using transthoracic aortic occlusion. *The Annals of thoracic surgery*, 63(5):1477–1479, 1997.

[30] Gordon Clapworthy, Marco Viceconti, Peter V Coveney, and Peter Kohl. The virtual physiological human: building a framework for computational biomedicine.

Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 366(1878):2975–2978, 2008.

[31] William J. Cook, William H. Cunningham, William R. Pulleyblank, and Alexander Schrijver. *Combinatorial optimization*. John Wiley & Sons, Inc., New York, NY, USA, 1998.

[32] T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. *IEEE Trans. Pattern Anal. Machine Intell.*, 23(6):681–685, 2001.

[33] T.F. Cootes, C.J. Taylor, D.H. Cooper, and J. Graham. Active shape models: Their training and application. *Computer Vision and Image Understanding*, 61(1):38–59, January 1995.

[34] Alain Cribier. Development of transcatheter aortic valve implantation (tavi): A 20-year odyssey. *Archives of cardiovascular diseases*, 105(3):146–152, 2012.

[35] A. Criminisi, D. Robertson, E. Konukoglu, J. Shotton, S. Pathak, S. White, and K. Siddiqui. Regression forests for efficient anatomy detection and localization in computed tomography scans. *Medical Image Analysis*, (0):–, 2013.

[36] Antonio Criminisi, Toby Sharp, and Andrew Blake. Geos: Geodesic image segmentation. In *Proc. European Conf. Computer Vision*, pages 99–112. Springer, 2008.

[37] David Cristinacce and Tim Cootes. Automatic feature localisation with constrained local models. *Pattern Recognition*, 41(10):3054 – 3067, 2008.

[38] Adnan S Dajani, Kathryn A Taubert, Walter Wilson, Ann F Bolger, Arnold Bayer, Patricia Ferrieri, Michael H Gewitz, Stanford T Shulman, Soraya Nouri, Jane W Newburger, et al. Prevention of bacterial endocarditis recommendations by the american heart association. *Circulation*, 96(1):358–366, 1997.

[39] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 886–893 vol. 1, 2005.

[40] Marleen de Bruijne, Bram van Ginneken, Max A Viergever, and Wiro J Niessen. Adapting active shape models for 3d segmentation of tubular structures in medical images. In *Information Processing in Medical Imaging*, pages 136–147. Springer, 2003.

[41] Carlos Domingo and Osamu Watanabe. Madaboost: A modification of adaboost. In *COLT*, pages 180–189. Citeseer, 2000.

[42] Lloyd-Jones Donald, Adams Robert, Carnethon Mercedes, De Simone Giovanni, Ferguson T. Bruce, Flegal Katherine, Ford Earl, Karen Furie, Go Alan, Greenlund Kurt, Haase Nancy, Hailpern Susan, Ho Michael, Howard Virginia, Kissela Brett,

Kittner Steven, Lackland Daniel, Lisabeth Lynda, Marelli Ariane, McDermott Mary, Meigs James, Mozaffarian Dariush, Nichol Graham, O'Donnell Christopher, Roger Veronique, Rosamond Wayne, Sacco Ralph, Sorlie Paul, Stafford Randall, Steinberger Julia, Thom Thomas, Wasserthiel-Smoller Sylvia, Wong Nathan, Wylie-Rosett Judith, and Hong Yuling. Heart disease and stroke statistics–2009 update: a report from the american heart association statistics committee and stroke statistics subcommittee. *Circulation*, 119(3), January 2009.

[43] D.R. Einstein, F. Del Pin, X. Jiao, A.P. Kuprat, J.P. Carson, K.S. Kunzelman, R.P. Cochran, J.M. Guccione, and M.B. Ratcliffe. Fluid-structure interactions of the mitral valve and left heart: Comprehensive strategies, past, present and future. *International Journal for Numerical Methods in Biomedical Engineering*, 26(3-4):348–380, 2010.

[44] P. Lang et al. Feature-based US to CT registration of the aortic root. In *Proceedings of SPIE Medical Imaging 2010*, 2010.

[45] Rebecca Fahrig, Robert Dixon, Thomas Payne, Richard L. Morin, Arundhuti Ganguly, and Norbert Strobel. Dose and image quality for a cone-beam c-arm ct system. *Medical Physics*, 33(12):4541–4550, 2006.

[46] Paul WM Fedak, Subodh Verma, Tirone E David, Richard L Leask, Richard D Weisel, and Jagdish Butany. Clinical and pathophysiological implications of a bicuspid aortic valve. *Circulation*, 106(8):900–904, 2002.

[47] Ted Feldman, Saibal Kar, Michael Rinaldi, Peter Fail, James Hermiller, Richard Smalling, Patrick L Whitlow, William Gray, Reginald Low, Howard C Herrmann, et al. Percutaneous mitral repair with the mitraclip systemsafety and midterm durability in the initial everest (endovascular valve edge-to-edge repair study) cohort. *Journal of the American College of Cardiology*, 54(8):686–694, 2009.

[48] James D Foley. *Computer graphics: Principles and practice, in C*, volume 12110. Addison-Wesley Professional, 1996.

[49] L. R. Ford and D. R. Fulkerson. *Flows in Networks*. Princeton University Press, 1962.

[50] W.T. Freeman and E.H. Adelson. The design and use of steerable filters. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 13(9):891 –906, sep 1991.

[51] Y. Freund and R. E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119–139, 1997.

[52] Jerome Friedman, Trevor Hastie, and Robert Tibshirani. Additive logistic regression: a statistical view of boosting. *Annals of Statistics*, 28:2000, 1998.

[53] Jerome Friedman, Trevor Hastie, and Robert Tibshirani. Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). *The annals of statistics*, 28(2):337–407, 2000.

[54] D. Fritz, D. Rinck, R. Dillmann, and M. Scheuring. Segmentation of the left and right cardiac ventricle using a combined bi-temporal statistical model. In *SPIE Medical Imaging*, pages 605–614, 2006.

[55] Valentin Fuster, R Wayne Alexander, Robert A O'Rourke, R Wayne Alexander, Robert A O'Rourke, Robert Roberts, Spencer B King, and Hein JJ Wellens. *Hurst's the Heart, 2-Vol Set*. McGraw-Hill Professional Publishing, 2000.

[56] Gang Gao, Graeme Penney, Nicolas Gogin, Pascal Cathier, Aruna Arujuna, Matt Wright, Dennis Caulfield, Aldo Rinaldi, Reza Razavi, and Kawal Rhode. Rapid image registration of three-dimensional transesophageal echocardiography and x-ray fluoroscopy for the guidance of cardiac interventions. In *Information Processing in Computer-Assisted Interventions*, pages 124–134. Springer, 2010.

[57] Gang Gao, Graeme Penney, Yingliang Ma, Nicolas Gogin, Pascal Cathier, Aruna Arujuna, Geraint Morton, Dennis Caulfield, Jaswinder Gill, C Aldo Rinaldi, et al. Registration of 3d trans-esophageal echocardiography to x-ray fluoroscopy using image-based probe tracking. *Medical image analysis*, 16(1):38–49, 2012.

[58] B Georgescu, XS Zhou, D. Comaniciu, and A. Gupta. Database-Guided Segmentation of Anatomical Structures with Complex Appearance. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 2, pages 429–436. IEEE, 2005.

[59] Ben Glocker, Johannes Feulner, Antonio Criminisi, David R Haynor, and Ender Konukoglu. Automatic localization and identification of vertebrae in arbitrary field-of-view ct scans. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012*, pages 590–598. Springer Berlin Heidelberg, 2012.

[60] C Gohlke-Bärwolf, J Acar, C Oakley, E Butchart, D Burckhardt, E Bodnar, R Hall, J-P Delahaye, D Horstkotte, R Kremer, et al. Guidelines for prevention of thromboembolic events in valvular heart disease. *European heart journal*, 16(10):1320–1330, 1995.

[61] Zehra Gölbasi, Özgül Uçar, Telat Keles, Ahmet Sahin, Kerim Çagli, Ahmet Çamsari, Erdem Diker, and Sinan Aydogdu. Increased levels of high sensitive c-reactive protein in patients with chronic rheumatic valve disease: evidence of ongoing inflammation. *European Journal of Heart Failure*, 4(5):593–595, 2002.

[62] Andrew V Goldberg and Robert E Tarjan. A new approach to the maximum flow problem. *Journal of the ACM (JACM)*, 35(4):921–940, 1988.

[63] C. Goodall. Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society. Series B (Methodological)*, 53(2):285–339, 1991.

[64] Leo Grady. Multilabel random walker image segmentation using prior models. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 763–770. IEEE, 2005.

[65] Sasa Grbic, Christian Gesell, R Lonasec, Matthias John, Jan Boese, Joachim Hornegger, Nassir Navab, and Dorin Cotnaniciu. Model-based fusion of ct and non-contrasted 3d c-arm ct: Application to transcatheter valve therapies. In *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*, pages 1192–1195. IEEE, 2012.

[66] Sasa Grbic, Razvan Ionasec, Tommaso Mansi, Bogdan Georgescu, Fernando Vegahiguera, Nassir Navab, and Dorin Comaniciu. Advanced Intervention Planning for Transcatheter Aortic Valve Implantation (TAVI) from CT using Volumetric Models. *2013 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2013.

[67] Sasa Grbić, Razvan Ionasec, Dime Vitanovski, Ingmar Voigt, Yang Wang, Bogdan Georgescu, Nassir Navab, and Dorin Comaniciu. Complete valvular heart apparatus model from 4D cardiac CT. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 13(Pt 1):218–26, January 2010.

[68] Sasa Grbic, Razvan Ionasec, Dime Vitanovski, Ingmar Voigt, Yang Wang, Bogdan Georgescu, Nassir Navab, and Dorin Comaniciu. Complete valvular heart apparatus model from 4D cardiac CT. *Medical image analysis*, 16(5):1003–14, July 2012.

[69] Saša Grbić, Razvan Ionasec, Yang Wang, Tommaso Mansi, Bogdan Georgescu, Matthias John, Jan Boese, Yefeng Zheng, Nassir Navab, and Dorin Comaniciu. Model-based fusion of multi-modal volumetric images: application to transcatheter valve procedures. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2011*, pages 219–226, 2011.

[70] Saša Grbic, Razvan I Ionasec, Dominik Zäuner, Yefeng Zheng, Bogdan Georgescu, and Dorin Comaniciu. Aortic valve and ascending aortic root modeling from 3d and 3d+ t ct. pages 76250H–76250H, 2010.

[71] Sasa Grbić, Tommaso Mansi, Razvan Ionasec, Bogdan Georgescu, Max Schoebinger, Nassir Navab, and Dorin Comaniciu. Image-Based Computational Models for TAVI Planning : From CT Images to Implant Deployment. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2013.

[72] Saša Grbic, Razvan I. Ionasec, Dominik Zäuner, Yefeng Zheng, Bogdan Georgescu, and Dorin Comaniciu. Aortic valve and ascending aortic root modeling from 3D and 3D+t CT. In Kenneth H. Wong and Michael I. Miga, editors, *SPIE Medical ...*, pages 76250H–76250H–8, March 2010.

[73] Martin Haensig, Lukas Lehmkuhl, Ardawan Julian Rastan, Joerg Kempfert, Chirojit Mukherjee, Matthias Gutberlet, David Michael Holzhey, and Friedrich Wilhelm Mohr. Aortic valve calcium scoring is a predictor of significant paravalvular aortic insufficiency in transapical-aortic valve implantation. *European Journal of Cardio-Thoracic Surgery*, 41(6):1234–1241, 2012.

[74] Joseph V Hajnal and Derek LG Hill. *Medical image registration*. CRC press, 2010.

[75] Joseph V Hajnal, Nadeem Saeed, Elaine J Soar, Angela Oatridge, Ian R Young, and Graeme M Bydder. A registration and interpolation procedure for subvoxel matching of serially acquired mr images. *Journal of computer assisted tomography*, 19(2):289–296, 1995.

[76] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*, volume 2. Cambridge Univ Press, 2000.

[77] Kentaro Hayashida, Erik Bouvier, Thierry Lefèvre, Bernard Chevalier, Thomas Hovasse, Mauro Romano, Philippe Garot, Yusuke Watanabe, Arnaud Farge, Patrick Donzeau-Gouge, et al. Transcatheter aortic valve implantation for patients with severe bicuspid aortic valve stenosis. *Circulation: Cardiovascular Interventions*, 6(3):284–291, 2013.

[78] Tobias Heimann and Hans-Peter Meinzer. Statistical shape models for 3D medical image segmentation: A review. *Medical image analysis*, 13(4):543, 2009.

[79] Tobias Heimann, Peter Mountney, Matthias John, and Razvan Ionasec. Learning without labeling: Domain adaptation for ultrasound transducer localization. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*, pages 544–551. Springer Berlin Heidelberg, 2013.

[80] Katrin Hemmann, Margarita Sirotina, Salvatore De Rosa, Joachim R. Ehrlich, Henrik Fox, Johannes Weber, Anton Moritz, Andreas M. Zeiher, Ilona Hofmann, Volker Schächinger, Mirko Doss, Horst Sievert, Stephan Fichtlscherer, and Ralf Lehmann. The sts score is the strongest predictor of long-term survival following transcatheter aortic valve implantation, whereas access route (transapical versus transfemoral) has no predictive value beyond the periprocedural phase. *Interactive CardioVascular and Thoracic Surgery*, 2013.

[81] John H Hipwell, Graeme P Penney, Tim C Cox, James V Byrne, and David J Hawkes. 2d-3d intensity based registration of dsa and mra-a comparison of similarity measures. In *Medical Image Computing and Computer Assisted Intervention MICCAI 2002*, pages 501–508. Springer, 2002.

[82] Julien IE Hoffman and Samuel Kaplan. The incidence of congenital heart disease. *Journal of the American College of Cardiology*, 39(12):1890–1900, 2002.

[83] Berthold K. P. Horn. Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America A*, 4(4):629–642, 1987.

[84] Berthold KP Horn. Closed-form solution of absolute orientation using unit quaternions. *JOSA A*, 4(4):629–642, 1987.

[85] Aapo Hyvärinen, Juha Karhunen, and Erkki Oja. *Independent Component Analysis*. John Wiley, Chichester, 2001.

[86] L. Ibanez, W. Schroeder, L. Ng, and J. Cates. *The ITK Software Guide*. Kitware, Inc., second edition, 2005.

[87] Razvan Ioan Ionasec, Ingmar Voigt, Bogdan Georgescu, Helene Houle, Joachim Hornegger, Nassir Navab, and Dorin Comaniciu. Personalized Modeling and Assessment of the Aortic-Mitral Coupling from 4D TEE and CT. In *MICCAI*, pages 767–775, Heidelberg, 2009.

[88] R.I. Ionasec, I. Voigt, B. Georgescu, Yang Wang, H. Houle, F. Vega-Higuera, N. Navab, and D. Comaniciu. Patient-specific modeling and quantification of the aortic and mitral valves from 4-d cardiac ct and tee. *Medical Imaging, IEEE Transactions on*, 29(9):1636 –1651, sept. 2010.

[89] Bernard Iung, Gabriel Baron, Eric G Butchart, François Delahaye, Christa Gohlke-Bärwolf, Olaf W Levang, Pilar Tornos, Jean-Louis Vanoverschelde, Frank Vermeer, Eric Boersma, et al. A prospective survey of patients with valvular heart disease in europe: The euro heart survey on valvular heart disease. *European heart journal*, 24(13):1231–1243, 2003.

[90] Tommi Jaakkola and David Haussler. Exploiting generative models in discriminative classifiers. *Advances in neural information processing systems*, pages 487–493, 1999.

[91] Alex Jablokow. National center for health statistics: National hospital discharge survey: Annual summaries with detailed diagnosis and procedure data. *Data on Health Resources Utilization*, 13, 2009.

[92] Mark Jenkinson and Stephen Smith. A global optimisation method for robust affine registration of brain images. *Medical image analysis*, 5(2):143–156, 2001.

[93] Daniel John, Lutz Buellesfeld, Seyrani Yuecel, Ralf Mueller, Georg Latsios, Harald Beucher, Ulrich Gerckens, and Eberhard Grube. Correlation of Device landing zone calcification and acute procedural success in patients undergoing transcatheter aortic valve implantations with the self-expanding CoreValve prosthesis. *JACC. Cardiovascular interventions*, 3(2):233–43, February 2010.

[94] Matthias John, Rui Liao, Yefeng Zheng, Alois Nöttling, Jan Boese, Uwe Kirschstein, Jörg Kempfert, and Thomas Walther. System to guide transcatheter aortic valve

implantations based on interventional c-arm ct imaging. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010*, pages 375–382. Springer, 2010.

[95] A Jordan. On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. *Advances in neural information processing systems*, 14:841, 2002.

[96] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *Int. J. Computer Vision*, 1(4):321–331, 1988.

[97] M.R. Kaus, J. von Berg, W. Niessen, and V. Pekar. Automated segmentation of the left ventricle in cardiac mri. In *MICCAI*, pages 432–439, 2003.

[98] Zeger F Knops, JB Antoine Maintz, Max A Viergever, and Josien PW Pluim. Normalized mutual information based registration using k-means clustering and shading correction. *Medical image analysis*, 10(3):432–439, 2006.

[99] Susheel K. Kodali, Mathew R. Williams, Craig R. Smith, Lars G. Svensson, John G. Webb, Raj R. Makkar, Gregory P. Fontana, Todd M. Dewey, Vinod H. Thourani, Augusto D. Pichard, Michael Fischbein, Wilson Y. Szeto, Scott Lim, Paul S. Greason, S. Chris Malaisrie, Pamela S. Douglas, Rebecca T. Hahn, Duolao Whisenant, Jodi J. Akin, William N. Anderson, and Martin B. Leon. Two-year outcomes after transcatheter or surgical aortic-valve replacement. *NEJM*, 366(18):1686, 2012.

[100] Vladimir Kolmogorov and Ramin Zabih. Computing visual correspondence with occlusions using graph cuts. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 2, pages 508–515. IEEE, 2001.

[101] Vladimir Kolmogorov and Ramin Zabin. What energy functions can be minimized via graph cuts? *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(2):147–159, 2004.

[102] Ender Konukoglu, Antonio Criminisi, Sayan Pathak, Duncan Robertson, Steve White, David Haynor, and Khan Siddiqui. Robust linear registration of ct images using random regression forests. In *SPIE Medical Imaging*, pages 79621X–79621X. International Society for Optics and Photonics, 2011.

[103] Ender Konukoglu, Ben Glocker, Darko Zikic, and Antonio Criminisi. Neighbourhood approximation using randomized forests. *Medical image analysis*, 2013.

[104] G. Krishnamurthy, A. Itoh, W. Bothe, J.C. Swanson, E. Kuhl, M. Karlsson, D. Craig Miller, and N.B. Ingels. Stress–strain behavior of mitral valve leaflets in the beating ovine heart. *Journal of biomechanics*, 42(12):1909–1916, 2009.

[105] Ayse Gul Kunt, Murat Kurtcephe, Mete Hidiroglu, Levent Cetin, Aslihan Kucuker, Vedat Bakuy, Ahmet Ruchan Akar, and Erol Sener. Comparison of original euroscore, euroscore ii and sts risk models in a turkish cardiac surgical cohort. *Interactive CardioVascular and Thoracic Surgery*, 16(5):625–629, 2013.

[106] KS Kunzelman and KP Cochran. Mechanical properties of basal and marginal mitral valve chordae tendineae. *ASAIO Journal*, 36(3):M405, 1990.

[107] KS Kunzelman, RP Cochran, C. Chuong, WS Ring, ED Verrier, and RD Eberhart. Finite element analysis of the mitral valve. *The Journal of heart valve disease*, 2(3):326, 1993.

[108] P. Lang, P. Seslija, M.W.A. Chu, D. Bainbridge, G.M. Guiraudon, D.L. Jones, and T.M. Peters. US-Fluoroscopy Registration for Transcatheter Aortic Valve Implantation. *IEEE Transactions on Biomedical Engineering*, 59(5):1444–1453, 2012.

[109] Emmanuel Lansac, Khee Hiang Lim, Yu Shomura, Wolfgang A. Goetz, Hou Sen Lim, Nolan T. Rice, Hashim Saber, and Carlos M.G. Duran. Dynamic balance of the aortomitral junction. *The Journal of Thoracic and Cardiovascular Surgery*, 123(5):911 – 918, 2002.

[110] Günter Lauritsch, Jan Boese, L Wigstrom, Herbert Kemeth, and Rebecca Fahrig. Towards cardiac c-arm computed tomography. *Medical Imaging, IEEE Transactions on*, 25(7):922–934, 2006.

[111] Alexander W Leber, Markus Kasel, Thomas Ischinger, Ulrich H Ebersberger, Diethmar Antoni, Martin Schmidt, Gotthard Riess, Vivian Renz, Armin Huber, Thomas Helmberger, et al. Aortic valve calcium score as a predictor for outcome after tavi using the corevalve revalving system. *International Journal of Cardiology*, 2011.

[112] Martin B. Leon, Craig R. Smith, Michael Mack, D. Craig Miller, Jeffrey W. Moses, Lars G. Svensson, E. Murat Tuzcu, John G. Webb, Gregory P. Fontana, Raj R. Makkar, David L. Brown, Peter C. Block, Robert A. Guyton, Augusto D. Pichard, Joseph E. Bavaria, Howard C. Herrmann, Pamela S. Douglas, John L. Petersen, Jodi J. Akin, William N. Anderson, Duolao Wang, and Stuart Pocock. Transcatheter aortic-valve implantation for aortic stenosis in patients who cannot undergo surgery. *New England Journal of Medicine*, 363(17):1597–1607, 2010.

[113] Rui Liao, Li Zhang, Ying Sun, Shun Miao, and Christophe Chefd'hotel. A review of recent advances in registration techniques applied to minimally invasive therapy. 2013.

[114] C. Lindner, S. Thiagarajah, J.M. Wilkinson, The arcOGEN Consortium , G.A. Wallis, and T.F. Cootes. Fully automatic segmentation of the proximal femur using random forest regression voting, 2013.

[115] Cristian Linte, Marcin Wierzbicki, John Moore, Stephen H. Little, Gerard M. Guiraudon, and Terry M. Peters. Towards subject-specific models of the dynamic heart for image-guided mitral valve surgery. In *MICCAI*, pages 94–101. 2007.

[116] Dong C Liu and Jorge Nocedal. On the limited memory BFGS method for large scale optimization. *Mathematical Programming*, 45(1-3):503–528, 1989.

[117] Harel Livyatan, Ziv Yaniv, and Leo Joskowicz. Gradient-based 2-d/3-d rigid registration of fluoroscopic x-ray to ct. *Medical Imaging, IEEE Transactions on*, 22(11):1395–1406, 2003.

[118] William E Lorensen and Harvey E Cline. Marching cubes: A high resolution 3d surface construction algorithm. In *ACM Siggraph Computer Graphics*, volume 21, pages 163–169. ACM, 1987.

[119] M. Lorenzo-Valdis, G. I. Sanchez-Ortiz, R. Mohiaddin, and D. Rueckert. Atlas-based segmentation and tracking of 3d cardiac mr images using non-rigid registration. In *MICCAI*, volume 2488, pages 642–650. 2002.

[120] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *Int. J. Computer Vision*, 60(2):91–110, 2004.

[121] Yongning Lu, Ying Sun, Rui Liao, and Sim Heng Ong. Registration of pre-operative ct and non-contrast-enhanced c-arm ct: an application to trans-catheter aortic valve implantation (tavi). In *Computer Vision–ACCV 2012*, pages 268–280. Springer, 2013.

[122] James T. Maddux, Onno Wink, John C. Messenger, Bertron M. Groves, Robert Liao, Jodi Strzelczyk, Shiuh-Yung Chen, and John D. Carroll. Randomized Study of the Safety and Clinical Utility of Rotational Angiography Versus Standard Angiography in the Diagnosis of Coronary Artery Disease. *Catheter. Cardiovasc. Interv.*, 62(2):167–174, 2004.

[123] Frederik Maes, Dirk Vandermeulen, and Paul Suetens. Medical image registration using mutual information. *Proceedings of the IEEE*, 91(10):1699–1722, 2003.

[124] George D Magoulas and Andriana Prentza. Machine learning in medical applications. In *Machine Learning and its applications*, pages 300–307. Springer, 2001.

[125] JB Maintz and Max A Viergever. A survey of medical image registration. *Medical image analysis*, 2(1):1–36, 1998.

[126] Timo Makela, Patrick Clarysse, Outi Sipila, Nicoleta Pauna, Quoc Cuong Pham, Toivo Katila, and Isabelle E Magnin. A review of cardiac image registration methods. *Medical Imaging, IEEE Transactions on*, 21(9):1011–1021, 2002.

[127] Tommaso Mansi, Stanley Durrleman, Boris Bernhardt, Maxime Sermesant, Hervé Delingette, Ingmar Voigt, Philipp Lurz, Andrew M Taylor, Julie Blanc, Younes Boudjemline, et al. A statistical model of right ventricle in tetralogy of fallot for prediction of remodelling and therapy planning. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009*, pages 214–221. Springer Berlin Heidelberg, 2009.

[128] Primoz Markelj, D Tomaževič, Bostjan Likar, and F Pernuš. A review of 3d/2d registration methods for image-guided interventions. *Medical image analysis*, 16(3):642–661, 2012.

[129] Barry J Maron. Hypertrophic cardiomyopathy. *JAMA: the journal of the American Medical Association*, 287(10):1308–1320, 2002.

[130] David Mattes, David R Haynor, Hubert Vesselle, Thomas K Lewellen, and William Eubank. Nonrigid multimodality image registration. In *Proceedings of SPIE Medical Imaging: Image Processing*, volume 4322, pages 1609–1620, July 2001.

[131] Tim McInerney and Demetri Terzopoulos. Deformable models in medical image analysis: a survey. *Medical Image Analysis*, 1(2):91 – 108, 1996.

[132] Patrick W McLaughlin, Vrinda Narayana, Marc Kessler, Daniel McShan, Sara Troyer, Lon Marsh, George Hixson, and Peter L Roberson. The use of mutual information in registration of ct and mri datasets post permanent implant. *Brachytherapy*, 3(2):61–70, 2004.

[133] Inc. Company Millennium Research Group, A Decision Resources. Us markets for heart valve devices 2012. 2012.

[134] Niloy J. Mitra, Natasha Gelfand, Helmut Pottmann, and Leonidas Guibas. Registration of Point Cloud Data from a Geometric Optimization Perspective. In *Proc. of Symp. on Geom. Proc.*, pages 22–31. ACM New York, 2004.

[135] Jan Modersitzki. *Numerical Methods for Image Registration (Numerical Mathematics and Scientific Computation)*. Oxford university press USA, 2004.

[136] Peter Mountney, Razvan Ionasec, Markus Kaizer, Sina Mamaghani, Wen Wu, Terrence Chen, Matthias John, Jan Boese, and Dorin Comaniciu. Ultrasound and fluoroscopic images fusion by autonomous ultrasound probe detection. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012*, pages 544–551. Springer Berlin Heidelberg, 2012.

[137] Keelin Murphy, Bram van Ginneken, Joseph M. Reinhardt, Sven Kabus, Kai Ding, Xiang Deng, Kunlin Cao, Kaifang Du, Gary E. Christensen, Vincent Garcia, Tom Vercauteren, Nicholas Ayache, Olivier Commowick, Grégoire Malandain, Ben Glocker, Nikos Paragios, Nassir Navab, Vladlena Gorbunova, Jon Sporring, Marleen de Bruijne, Xiao Han, Mattias P. Heinrich, Julia A. Schnabel, Mark Jenkinson, Cristian

Lorenz, Marc Modat, Jamie McClelland, Sébastien Ourselin, Sascha E. A. Muenzing, Max A. Viergever, Dante De Nigris, D. Louis Collins, Tal Arbel, Marta Peroni, Rui Li, Gregory C. Sharp, Alexander Schmidt-Richberg, Jan Ehrhardt, René Werner, Dirk Smeets, Dirk Loeckx, Gang Song, Nicholas J. Tustison, Brian B. Avants, James C. Gee, Marius Staring, Stefan Klein, Berend C. Stoel, Martin Urschler, Manuel Werlberger, Jef Vandemeulebroucke, Simon Rit, David Sarrut, and Josien P. W. Pluim. Evaluation of registration methods on thoracic ct: The empire10 challenge. *IEEE Trans. Med. Imaging*, 30(11):1901–1920, 2011.

[138] Diab Mutlak, D. Aronson, J. Lessick, S. Reisner, S. Dabbah, and Y. Agmon. Functional tricuspid regurgitation in patients with pulmonary hypertension. *CHEST*, pages 115–121, 2009.

[139] Andriy Myronenko and Xubo Song. Point set registration: Coherent point drift. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(12):2262–2275, 2010.

[140] Andriy Myronenko, Xubo Song, and Miguel A Carreira-Perpinán. Non-rigid point set registration: Coherent point drift. In *Advances in Neural Information Processing Systems*, pages 1009–1016, 2006.

[141] Thomas Netsch, Peter Rosch, Arianne van Muiswinkel, and Jürgen Weese. Towards real-time multi-modality 3-d medical image registration. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 1, pages 718–725. IEEE, 2001.

[142] Dominik Neumann, Sasa Grbic, Matthias John, Nassir Navab, Joachim Hornegger, and Razvan Ioan Ionasec. Robust model-based 3d/3d fusion using sparse matching for minimally invasive surgery. In *Medical Image Computing and Computer Assisted Intervention*, Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2013.

[143] Arnold CT Ng, Victoria Delgado, Frank van der Kley, Miriam Shanks, Nico RL van de Veire, Matteo Bertini, Gaetano Nucifora, Rutger J van Bommel, Laurens F Tops, Arend de Weger, et al. Comparison of aortic root dimensions and geometries before and after transcatheter aortic valve implantation by 2-and 3-dimensional transesophageal echocardiography and multislice computed tomography. *Circulation: Cardiovascular Imaging*, 3(1):94–102, 2010.

[144] Jim Oeppen and James W Vaupel. Broken limits to life expectancy. *Science*, 296(5570):1029–1031, 2002.

[145] M. Oren, C. Papageorgiou, P. Sinha, E. Osuna, and T. Poggio. Pedestrian detection using wavelet templates. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages 193–199, 1997.

[146] Robert C. Orth, Michael J. Wallace, and Michael D. Kuo. C-arm cone-beam ct: General principles and technical considerations for use in interventional radiology. *Journal of Vascular and Interventional Radiology*, 19(6):814 – 820, 2008.

[147] C. P. Papageorgiou, M. Oren, and T. Poggio. A general framework for object detection. In *ICCV '98: Proceedings of the Sixth International Conference on Computer Vision*, page 555, Washington, DC, USA, 1998. IEEE Computer Society.

[148] Olivier Pauly, Ben Glocker, Antonio Criminisi, Diana Mateus, Axel Martinez Möller, Stephan Nekolla, and Nassir Navab. Fast multiple organ detection and localization in whole-body mr dixon sequences. In *Proceedings of the 14th international conference on Medical image computing and computer-assisted intervention - Volume Part III*, MICCAI'11, pages 239–247, Berlin, Heidelberg, 2011. Springer-Verlag.

[149] Graeme P Penney, Jürgen Weese, John A Little, Paul Desmedt, Derek LG Hill, et al. A comparison of similarity measures for use in 2-d-3-d medical image registration. *Medical Imaging, IEEE Transactions on*, 17(4):586–595, 1998.

[150] Terry M Peters. Image-guidance for surgical procedures. *Physics in medicine and biology*, 51(14):R505, 2006.

[151] Josien PW Pluim, JB Antoine Maintz, and Max A Viergever. Mutual-information-based registration of medical images: a survey. *Medical Imaging, IEEE Transactions on*, 22(8):986–1004, 2003.

[152] William H Press. *Numerical recipes in Fortran 77: the art of scientific computing*, volume 1. Cambridge university press, 1992.

[153] V. Prot, R. Haaverstad, and B. Skallerud. Finite element analysis of the mitral apparatus: annulus shape effect and chordal force distribution. *Biomechanics and modeling in mechanobiology*, 8(1):43–55, 2009.

[154] V. Prot and B. Skallerud. Nonlinear solid finite element analysis of mitral valves with heterogeneous leaflet layers. *Computational Mechanics*, 43(3):353–368, 2009.

[155] William C Roberts. The congenitally bicuspid aortic valve: a study of 85 autopsy cases. *The American journal of cardiology*, 26(1):72–83, 1970.

[156] F. Roques, S.A.M. Nashef, P. Michel, E. Gauducheau, C. de Vincentiis, E. Baudet, J. Cortina, M. David, A. Faichney, F. Gavrielle, E. Gams, A. Harjula, M.T. Jones, P. Pinna Pintor, R. Salamon, and L. Thulin. Risk factors and outcome in european cardiac surgery: analysis of the euroscore multinational database of 19030 patients. *European Journal of Cardio-Thoracic Surgery*, 15(6):816–823, 1999.

[157] Raphael Rosenhek, Bernard Iung, Pilar Tornos, Manuel J. Antunes, Bernard D. Prendergast, Catherine M. Otto, Arie Pieter Kappetein, Janina Stepinska, Jens J. Kaden,

Christoph K. Naber, Esmeray Acartürk, and Christa Gohlke-Bärwolf. Esc working group on valvular heart disease position paper: assessing the risk of interventions in patients with valvular heart disease. *European Heart Journal*, 2011.

[158] Mikaël Rousson, Nikos Paragios, and Rachid Deriche. Implicit active shape models for 3d segmentation in mr imaging. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2004*, pages 209–216. Springer, 2004.

[159] M.R. Sabuncu, B.T.T. Yeo, K. Van Leemput, B. Fischl, and P. Golland. A generative model for image segmentation based on label fusion. *Medical Imaging, IEEE Transactions on*, 29(10):1714–1729, 2010.

[160] JasonM. Saragih, Simon Lucey, and JeffreyF. Cohn. Deformable model fitting by regularized landmark mean-shift. *International Journal of Computer Vision*, 91(2):200–215, 2011.

[161] David Sarrut and Sébastien Clippe. Geometrical transformation approximation for 2d/3d intensity-based registration of portal images and ct scan. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2001*, pages 532–540. Springer, 2001.

[162] R. E. Schapire, Y. Freund, P. Bartlett, and W.S. Lee. Boosting the margin: A new explanation for the effectiveness of voting methods. *The Annals of Statistics*, 26(5):1651–1686, 1998.

[163] S. Schievano, K. Kunzelman, M.A. Nicosia, R.P. Cochran, D.R. Einstein, S. Khambadkone, and P. Bonhoeffer. Percutaneous mitral valve dilatation: Single balloon versus double balloon. a finite element study. *Journal of Heart Valve Disease*, 18(1):28–34, 2009.

[164] R. Schneider, W. Burke, G. Marx, P. del Nido, and R. Howe. Modeling mitral valve leaflets from three-dimensional ultrasound. In *Functional Imaging and Modeling of the Heart–FIMH 2011*, LNCS, pages 215–222. Springer, 2011.

[165] R. Schneider, N. Tenenholtz, D. Perrin, G. Marx, P. del Nido, and R. Howe. Patient-specific mitral leaflet segmentation from 4D ultrasound. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2011*, LNCS, pages 520–527. Springer, 2011.

[166] R.J. Schneider, D.P. Perrin, N.V. Vasilyev, G.R. Marx, P.J. del Nido, and R.D. Howe. Mitral annulus segmentation from 3d ultrasound using graph cuts. *Medical Imaging, IEEE Transactions on*, 29(9):1676 –1687, sept. 2010.

[167] R.J. Schneider, D.P. Perrin, N.V. Vasilyev, G.R. Marx, P.J. del Nido, and R.D. Howe. Real-time image-based rigid registration of three-dimensional ultrasound. *Medical Image Analysis*, 2011.

[168] P. Schoenhagen, S. S. Halliburton, A. E. Stillman, and R. D. White. CT of the heart: Principles, advances, clinical uses. *Cleveland Clinic Journal of Medicine*, 72(2):127–138, 2005.

[169] CJ Schultz, AD Moelker, A Tzikas, RJ Geuns, PJ Feyter, and Serruys PW. Cardiac ct: necessary for precise sizing for transcatheter aortic implantation. *Eurointervention*, pages 6–13.

[170] Yoshiaki Shirai. *Three-dimensional computer vision*. Springer-verlag Berlin, 1987.

[171] L.H. Staib and J.S. Duncan. Boundary finding with parametrically deformable models. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 14(11):1061–1075, 1992.

[172] M. Stevanella, G. Krishnamurthy, E. Votta, J.C. Swanson, A. Redaelli, and N.B. Ingels. Mitral leaflet modeling: Importance of in vivo shape and material properties. *Journal of biomechanics*, 44(12):2229–2235, 2011.

[173] Marco Stevanella, Francesco Maffessanti, Carlo Conti, Emiliano Votta, Alice Arnoldi, Massimo Lombardi, Oberdan Parodi, Enrico Caiani, and Alberto Redaelli. Mitral valve patient-specific finite element modeling from cardiac mri: Application to an annuloplasty procedure. *Cardiovascular Engineering and Technology*, pages 1–11, 2011. 10.1007/s13239-010-0032-4.

[174] Colin Studholme, Derek LG Hill, and David J Hawkes. An overlap invariant entropy measure of 3d medical image alignment. *Pattern recognition*, 32(1):71–86, 1999.

[175] M-Pilar Tornos, G Permanyer-Miralda, Montserrat Olona, Miguel Gil, Enrique Galve, Benito Almirante, and J Soler-Soler. Long-term complications of native valve infective endocarditis in non-addictsa 15-year follow-up study. *Annals of internal medicine*, 117(7):567–572, 1992.

[176] A. Torralba, K.P. Murphy, and W.T. Freeman. Sharing features: efficient boosting procedures for multiclass object detection. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, volume 2, pages II–762–II–769 Vol.2, 2004.

[177] Wendy Tsang, Massimiliano Meineri, Rebecca T Hahn, Federico Veronesi, Atman P Shah, Mark Osten, Sandeep Nathan, Mark Russo, Roberto M Lang, and Eric M Horlick. A three-dimensional echocardiographic study on aortic–mitral coupling in transcatheter aortic valve replacement. *European Heart Journal–Cardiovascular Imaging*, 2013.

[178] Z. Tu. Probabilistic boosting-tree: Learning discriminative methods for classification, recognition, and clustering. In *International Conference on Computer Vision*, pages 1589–1596, 2005.

[179] Zhuowen Tu. Probabilistic boosting-tree: learning discriminative models for classification, recognition, and clustering. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, volume 2, pages 1589 –1596 Vol. 2, oct. 2005.

[180] Zhuowen Tu. Auto-context and its application to high-level vision tasks. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008.

[181] S Tzamtzis, J Viquerat, J Yap, MJ Mullen, and G Burriesci. Numerical analysis of the radial force produced by the Medtronic-CoreValve and Edwards-SAPIEN after transcatheter aortic valve implantation (TAVI). *Med. Eng. & Physics*, 2012.

[182] Shinji Umeyama. Least-squares estimation of transformation parameters between two point patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 13(4):376–380, 1991.

[183] Martin Urschler, Joachim Bauer, Hendrik Ditt, and Horst Bischof. Sift and shape context for feature-based nonlinear registration of thoracic ct images. In *Computer Vision Approaches to Medical Image Analysis*, pages 73–84. Springer Berlin Heidelberg, 2006.

[184] Alec Vahanian, Helmut Baumgartner, Jeroen Bax, Eric Butchart, Robert Dion, Gerasimos Filippatos, Frank Flachskampf, Roger Hall, Bernard Iung, Jaroslaw Kasprzak, et al. Guidelines on the management of valvular heart disease the task force on the management of valvular heart disease of the european society of cardiology. *European heart journal*, 28(2):230–268, 2007.

[185] Bram Van Ginneken, Alejandro F Frangi, Joes J Staal, Bart M ter Haar Romeny, and Max A Viergever. Active shape model segmentation with optimal features. *medical Imaging, IEEE Transactions on*, 21(8):924–933, 2002.

[186] Eva M. van Rikxoort, Ivana Isgum, Yulia Arzhaeva, Marius Staring, Stefan Klein, Max A. Viergever, Josien P.W. Pluim, and Bram van Ginneken. Adaptive local multi-atlas segmentation: Application to the heart and the caudate nucleus. *Medical Image Analysis*, 14(1):39 – 49, 2010.

[187] F Vanden-Eynden, D Bouchard, I El-Hamamsy, A Butnaru, P Demers, M Carrier, LP Perrault, JC Tardif, and M Pellerin. Effect of aortic valve replacement for aortic stenosis on severity of mitral regurgitation. *Ann Thorac Surg*, 83:1279–84, 2007.

[188] F. Veronesi, C. Corsi, L. Sugeng, V. Mor-Avi, E.G. Caiani, L. Weinert, C. Lamberti, and R.M. Lang. A study of functional anatomy of aortic-mitral valve coupling using 3D matrix transesophageal echocardiography. *Circ Cardiovasc Imaging*, 2(1):24–31, 2009.

[189] F Vinals, F Heredia, and A Giuliano. The role of the three vessels and trachea view (3vt) in the diagnosis of congenital heart defects. *Ultrasound in obstetrics & gynecology*, 22(4):358–367, 2003.

[190] P. Viola and M. Jones. Robust real-time face detection. *International Journal of Computer Vision*, 57(2):137–154, 2003.

[191] Paul Viola and William M Wells III. Alignment by maximization of mutual information. *International journal of computer vision*, 24(2):137–154, 1997.

[192] J. von Berg and C. Lorenz. Multi-surface cardiac modelling, segmentation, and tracking. In *Proc. Functional Imaging and Modeling of the Heart*, pages 1–11, 2005.

[193] E. Votta, E. Caiani, F. Veronesi, M. Soncini, F.M. Montevecchi, and A. Redaelli. Mitral valve finite-element modelling from ultrasound data: a pilot study for a new approach to understand mitral function and clinical scenarios. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 366(1879):3411, 2008.

[194] Emiliano Votta, Trung Bao Le, Marco Stevanella, Laura Fusini, Enrico G Caiani, Alberto Redaelli, and Fotis Sotiropoulos. Toward patient-specific simulations of cardiac valves: State-of-the-art and future directions. *Journal of biomechanics*, 2012.

[195] I. Waechter, R. Kneser, G. Korosoglou, J. Peters, N. Bakker, R. Boomen, and J. Weese. Patient specific models for planning and guidance of minimally invasive aortic valve implantation. In Tianzi Jiang, Nassir Navab, Josien Pluim, and Max Viergever, editors, *Medical Image Computing and Computer Assisted Intervention*, volume 6361 of *Lecture Notes in Computer Science*, pages 526–533. Springer Berlin / Heidelberg, 2010.

[196] Qian Wang, Eric Sirois, and Wei Sun. Patient-specific modeling of biomechanical interaction in transcatheter aortic valve deployment. *Journal of biomechanics*, 2012.

[197] Manfred K Warmuth, Karen A Glocer, and SVN Vishwanathan. Entropy regularized lpboost. In *Algorithmic Learning Theory*, pages 256–271. Springer, 2008.

[198] Manfred K Warmuth, Jun Liao, and Gunnar Rätsch. Totally corrective boosting algorithms that maximize the margin. In *Proceedings of the 23rd international conference on Machine learning*, pages 1001–1008. ACM, 2006.

[199] Wolfgang Wein, Barbara Röper, and Nassir Navab. Automatic registration and fusion of ultrasound with ct for radiotherapy. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2005*, pages 303–311. Springer, 2005.

[200] William M Wells, Paul Viola, Hideki Atsumi, Shin Nakajima, and Ron Kikinis. Multimodal volume registration by maximization of mutual information. *Medical image analysis*, 1(1):35–51, 1996.

[201] L. Yang, B. Georgescu, Y.F. Zheng, P. Meer, and D. Comaniciu. 3d ultrasound tracking of the left ventricle using one-step forward prediction and data fusion of collaborative trackers. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2008.

[202] Li Zhang, Christophe Chefd'hotel, Vincent Ordy, Jie Zheng, Xiang Deng, and Benjamin Odry. A knowledge-driven quasi-global registration of thoracic-abdominal CT and CBCT for image-guided interventions. In *Proceedings of SPIE Medical Imaging 2013: Image-Guided Procedures, Robotic Interventions, and Modeling*, volume 8671, March 2013.

[203] Zhengyou Zhang. Iterative point matching for registration of free-form curves and surfaces. *International journal of computer vision*, 13(2):119–152, 1994.

[204] Y. Zheng, B. Barbu, Georgescu, M. Scheuering, and D. Comaniciu. Fast automatic heart chamber segmentation from 3d ct data using marginal space learning and steerable features. 2007.

[205] Yefeng Zheng, A. Barbu, B. Georgescu, M. Scheuering, and D. Comaniciu. Four-chamber heart modeling and automatic segmentation for 3-d cardiac ct volumes using marginal space learning and steerable features. *Medical Imaging, IEEE Transactions on*, 27(11):1668 –1681, nov. 2008.

[206] Yefeng Zheng, Matthias John, Rui Liao, Jan Boese, Uwe Kirschstein, Bogdan Georgescu, S Kevin Zhou, Joerg Kempfert, Thomas Walther, Gernot Brockmann, and et al. Automatic aorta segmentation and valve landmark detection in c-arm ct: application to aortic valve implantation. *Medical Image Computing and Computer-Assisted Intervention*, 13(Pt 1):476–483, 2010.

[207] Yefeng Zheng, Fernando Vega-Higuera, Shaohua Kevin Zhou, and Dorin Comaniciu. Fast and Automatic Heart Isolation in 3D CT Volumes: Optimal Shape Initialization. In *MLMI*, volume 6357 of *LNCS*, pages 84–91. Springer Berlin Heidelberg, 2010.

[208] S.K. Zhou, Jie Shao, Bogdan Georgescu, and Dorin Comaniciu. Boostmotion: Boosting a discriminative similarity function for motion estimation. In *CVPR*, pages 1761–1768, 2006.

[209] Yun Zhu, X. Papademetris, A.J. Sinusas, and J.S. Duncan. Segmentation of the left ventricle from cardiac mr images using a subject-specific dynamical model. *Medical Imaging, IEEE Transactions on*, 29(3):669–687, 2010.

[210] Yun Zhu, Xenophon Papademetris, Albert Sinusas, and James Duncan. A dynamical shape prior for lv segmentation from rt3d echocardiography. In Guang-Zhong Yang, David Hawkes, Daniel Rueckert, Alison Noble, and Chris Taylor, editors, *Medical*

Image Computing and Computer Assisted Intervention, volume 5761 of *Lecture Notes in Computer Science*, pages 206–213. Springer Berlin / Heidelberg, 2009.

[211] Xiahai Zhuang, K.S. Rhode, R.S. Razavi, D.J. Hawkes, and S. Ourselin. A registration-based propagation framework for automatic whole heart segmentation of cardiac mri. *Medical Imaging, IEEE Transactions on*, 29(9):1612 –1625, sept. 2010.

[212] Lilla Zollei and William Wells. Multi-modal image registration using dirichlet-encoded prior information. In *Int. Workshop on Biomed. Image Registration*, pages 34–42. 2006.