‘In science, there is only physics; all the rest is stamp collecting.’

Ernest Rutherford (1871–1937)

Summary and Conclusions
9.1 Summary

"Aging of the population will undoubtedly result in a concomitant increase in the incidence of chronic diseases, including coronary artery disease, heart failure, and stroke" [9]. Due to the increasing prevalence of cardiac and cardiovascular diseases, an increasing number of diagnostic assessments and interventions will be performed in the future. Already, the amount of diagnostic assessments is increasing rapidly, and as a consequence hospitals perform an increasing number of imaging operations. This yields a huge amount of image data to be investigated, which consequently cannot be done manually or visually by the physicians any more, who need an increasing portion of their working time for interaction with patients and their treatment. It is therefore an obvious conclusion that assessment and diagnosis based on imaging data has to be automated to the maximal extent possible. This is the purpose and the justification of the research that has been presented in this thesis. In Chapter 1, the development of an automatic cardiac segmentation method was formulated as the main goal of this research. The requirements for the desired method set in Chapter 1 were:

- it should treat segmentation in an intrinsically three-dimensional manner and exploit 3D spatial continuity,
- it should be applicable to 3D cardiac image data, from different modalities or acquisition protocols without retraining of the statistical model,
- it should be able to segment image data irrespective of the orientation of its constituting image slices,
- it should be able to segment a data set using only a few images, either parallel or possibly with various orientations.

Below, a summary of the achievements presented in Chapters 2-8 will show that the developed methods perform equally well or better than methods from literature, while all requirements listed above are met.

In Chapter 2, a feasibility study was performed to assess whether 3D-ASMs are suitable for automatic segmentation of the cardiac left ventricle from MSCT image data. It was shown that the 3D active shape model that was developed based on shape information extracted from cardiac MRI data sets could be applied to cardiac MSCT data without changing anything to the statistical shape part. Because we omitted (statistical) gray level information from the training stage of our model, it appeared to be applicable to data from another modality than that used in the shape training stage. Edge information was deduced using extraction of pixel strips of one pixel wide and sufficiently long to capture the tissue transitions. Edge positions were obtained by convolution of the pixel strips with a gradient filter. The model converged well and produced visually usable contours for short-axis data sets (see Fig. 2.5). An initial application to axial data sets was also shown as a proof-of-concept.

Quantitative evaluation from the short-axis data sets however, showed that ED LV volume (blood pool plus myocardium) was systematically overestimated by the model, whereas ED blood pool is systematically underestimated (see Fig. 2.6). The fact that the apex was not included in the model and non-optimal parameter settings (for edge detection and model relaxation) may be the cause for these systematic errors observed with this model. On the other hand, excellent correlation factors (R=0.99) were observed for both the LV blood pool volume and epicardial volume with respect to their manual counterparts.

We have proved that a comprehensive representation of the heart shape can be achieved by building a statistical model utilizing the data acquired with non-invasive modalities. In addition, we have shown that the invariance of the organ shape with respect
to imaging with another (invasive) modality can be used for recognition purposes and automated contour detection.

In Chapter 3, a visual improvement in segmentation results with respect to the results from Chapter 2 was achieved by replacement of the filter based update mechanism with a Fuzzy Inference System in the edge detection part of the ASM. Thus, a more robust scheme for generating point update information was introduced. The FIS proved to be a more reliable system for identifying positions for point updates based on Fuzzy C-means classification of surrounding tissues. Furthermore, higher spatial coherence of update information extracted from patches was achieved using the FIS. The model was applied to a limited test data set of six patients, and consequently no extensive statistical evaluation was performed. A visual comparison of the results from Chapter 3 with the method used in Chapter 2 showed clear improvements in the final contour delineations (see Figs. 3.6 and 3.7). From quantitative results, the FIS-based method proved to be more accurate in blood pool volume and more precise in total LV volume.

In Chapter 4, the Fuzzy Inference System that was used in the previous chapter was elaborately described. The FIS involved classification of pixel gray values into several tissue classes using Fuzzy C-means clustering. Model generation, the use of the model mesh, and model matching were covered in detail. Parameters steering model behavior were presented in Table 4.1. The FIS-based 3D-ASM was applied to both MSCT and MRI data sets (25 patient data sets for MSCT and 15 healthy volunteer data sets for MRI). The results were quantitatively evaluated by assessment of point-to-point distances to manual segmentations, clinical contour quality, performing volume regression analysis and Bland-Altman analysis with respect to endocardial and epicardial volumes. The results proved that with a minimal effort of adjusting tunable parameters, the FIS-based 3D-ASM could be adapted for application to MRI or MSCT. Still LV volumes were systematically slightly overestimated and blood pool volumes underestimated (see Fig. 4.11), but clinically acceptable segmentation results for both modalities could be generated. The experimental results also confirmed the initial hypothesis, that no adaptation is required in the training stage of the model, relieving from the necessity of using invasive or toxic modalities for building models for different imaging modalities. Healthy volunteers will not likely consent nor be allowed to be scanned with a CT scanner just to collect a large training data set for a CT-based statistical shape model. With the availability of the 3D-ASM presented in Chapter 4, this is no longer necessary, as a statistical shape model derived from MRI training data also suffices.

With this achievement, the first two points of the goals of this research in Section 1.3 have been realized. Feasibility of the third point was demonstrated with an example (see Fig. 2.5(b)).

In Chapter 5, a grid-based optimization approach was presented, to achieve optimal settings for parameters that determine model behavior and performance in combination with application of the model to MRI data sets. In this chapter a new definition of the model mesh was used, now also including the apex. The mesh was constructed by automatic landmarking of the mean shape of a LV-atlas constructed with non-rigid registration of manually labeled LV-volumes. Segmentation performances of the model in both ED and ES, before and after grid-based parameter optimization were compared using point-to-surfaces distances from the model segmentations to manual segmentation, which served as the gold standard. These distances were calculated from points on the final model state to the closest locations on the manually segmented surface.
Optimal parameter settings are presented in Table 5.2. The evaluation data set used in this study consisted of 30 subjects: 15 healthy volunteers and 15 patients. Table 5.3 presents the final segmentation performance achieved after parameter optimization and shows the percentages (on the order of 13-28%) of segmentation performance improvement with respect to previously used ad hoc settings for the parameters. From this chapter, we conclude that in the very active and evolving field of medical image analysis grid-technology has become a real necessity. Grid technology solutions considerably shorten execution times of exhaustive searches and large-scale image processing, effectively enabling the sharing of computing resources between institutions. Finally, scientific progress and derived clinical applications greatly benefit from the scalability and computational power of the latest grid computing approaches.

In Chapter 6 three different ways of distributing landmarks over the surfaces of our model have been studied and evaluated with respect to their segmentation performances. The point correspondence methods are

- the automatic landmarking method in combination with the cardiac non-rigid registration-based atlas (also used in Chapter 5: almk). Landmarks are propagated to the individual subjects.
- the point distribution method used in Chapters 2, 3 and 4, with a regular grid derived by equiangular sampling, in every single subject, of the manually derived myocardial borders (rss).
- a hybrid point distribution, i.e., a mixture of the two other methods, achieved by equiangular resampling of the mean shape of the non-rigid registration-based atlas and propagation of the landmarks to the individual subjects (rsa).

All three methods were applied to the same test population as used in Chapter 5 and the segmentation performance was also evaluated in exactly the same manner.

With respect to model compactness, generalization, and specificity, not one model clearly outperformed the others for all these properties. For compactness, generalization and specificity, the rss correspondence definition seems to either close the gap with almk going from the 2-chamber model to the 1-chamber model, or increase its lead (which was marginal with respect to generalization ability). This leads to the hypothesis that the almk model definition is better for the 2-chamber model and that the rss correspondence definition is suitable for shapes as simple as the (circular) LV in the 1-chamber model.

It has to be noted that the almk model does not only differ from the other two models with respect to point correspondence. The fact that almk has either one or two closed apices (for the 1-chamber and 2-chamber models respectively) and the other models do not, may also cause differences with respect to compactness, generalization and specificity.

As far as segmentation performance is concerned, it appeared impossible to clearly distinguish between the 2-chamber models. Overall segmentation performance was hampered by confounding image clues in the right ventricle, preventing linked model structures from proper deformation (i.e., segmentation). Tests of the 1-chamber model on pre-segmented data indicated that both the rsa and almk models performed better than the rss model. Further, the almk and rsa models improved with allowing more modes of variation, whereas rss hardly shows any difference in performance. This is consistent with the observation that the rss model captures much of its total variance in the first few modes. With respect to segmentation performance on clinical data, almk seemed to be slightly better than rsa, followed by rss. These differences, however, are not significant.
In Chapter 7, a method (coined SPASM: SParse data Active Shape Model) for application of our 3D-ASM to sparse data sets was developed. Although the initial model was trained using densely acquired imaging data, SPASM offers a solution during the matching stage to cope with the absence of image information due to sparse sampling of cardiac regions. A novel update scheme was proposed based on propagation weighted with a Gaussian kernels. This allowed distribution of model update information from locations where image data is present to regions where this information is absent. Absence of image information was due to sparse sampling of the cardiac region, i.e. sparse scanning of the thorax during image acquisition. Propagation was stopped at a cut-off distance from the update source measured along the model surface. Using grid technology, optimal propagation settings were derived in the different data configurations, acquired with different MRI acquisition protocols, for which the propagation scheme was tested. For these tests the almk model from Chapters 5 and 6 was used. Optimal propagation settings for the different data configurations can be found in Table 7.1. Final segmentation results (see Tab. 7.3) were compared to results from literature and to interobserver variation with respect to manual segmentation. SPASM's performance on sparse data sets was comparable to or better than results reported in literature, but slightly worse than interobserver variation for manual segmentation. However, SPASM was applied to four image planes whereas the other methods used 8-12 planes. Moreover, SPASM was applied to arbitrarily oriented image planes, not depending on plane orientations present in the training data set, whereas the methods compared to all require equal data sampling density and orientation as present in their respective training sets. Sensitivity of SPASM to model placement during initialization was tested using grid technology by perturbation of the initial model position with steps of 10 mm within plus and minus 20 mm from the previously chosen initial position. Perturbations were added in the directions of the RV-to-LV axis, in the AP-direction and in the direction of the long-axis. Observed sensitivities to initial model placement show that initialization of the model is important, but within a range of 15-20 mm in all directions effects in final segmentations are minimal (see Figs. 7.11-7.13).

The incorporation of an update propagation scheme and a Fuzzy Inference System enabled application of SPASM to multi-protocol cardiac sparse data sets with a segmentation performance that is better than or comparable to other 3D model-based segmentation methods operating on a full data set with parallel image planes. Because SPASM does not include a statistical gray level model, it is applicable to data sets from different MRI acquisition protocols without adaptations to the statistical framework or the update determination scheme (FIS). With this achievement, all goals set out in Section 1.3 have been realized. Not only application to arbitrarily oriented image planes proved feasible, but segmentation performance also competes with that of other models presented in literature that need many more image slices, and do not allow freedom of orientation or sparsity.

Finally, in Chapter 8, the potential of SPASM for efficient reconstruction of LV surfaces and accurate volume quantification from limited information from sparsely acquired in-vivo imaging data with different spatial orientation was evaluated. SPASM was applied to radially oriented long-axis data sets (RAD), short-axis data sets (SA) and to a data configuration that consisted of a 2-chamber view, a 4-chamber view, and two short-axis views (multi-view, MV) (see Fig. 8.1). Subsets from the RAD data set encompassing two image slices and a range of sub sets from the SA data set were tested to assess at which data sparsity SPASM can still be applied without achieving statistically significant different (i.e., worse) segmentation results with respect to application
to a full data set. Performance was evaluated using point-to-surface distances and volumes of the blood pool and of the complete left ventricle (blood pool plus myocardium). Results showed that on the MV data set (with four planes) and the SA-6 data set (SA with 6 planes) SPASM does not perform significantly different than when applied to the full SA data set (11 planes). Systematic differences in final model volumes with respect to manual segmentation were observed due to the inclusion of the apex in the almk model, while the apex was not present in the manual segmentation. Based on this fact, the point-to-surface errors may be smaller when they are corrected for this.

9.2 Conclusions and future work

From everything presented in this thesis, we conclude that all goals stated in Section 1.3 have been realized. The segmentation results achieved are comparable to or better than those achieved with much more complex models, embedding substantially more knowledge. Applicability of the 3D-ASM and SPASM to MSCT, MRI, and within MRI to different acquisition protocols has been shown. In the future however, applicability to 3D ultrasound should be investigated.

The models presented were all built from and applied to cardiac imagery. However, the choices made in the development of the models allow them to be applied to other organs in the human body with only minor adjustments. The employment of a tissue classification scheme allows application to images from multiple modalities. Capturing only shape knowledge and removing gray level training led to independence with respect to imaging plane orientations. The image information update propagation scheme in SPASM allows application of the model to either more sparsely or more densely sampled organs. These achievements together suggest that the models are equally well applicable to imagery from other organs, provided that tissues in and around these organs are well classifiable and provided that the organ shapes are well characterizable with a suitable point correspondence. The kidneys, the liver and the lungs, e.g., would be perfect candidates to try to apply the models developed here to.

In this work, application has focused mainly on the LV, and little effort has been put into application to multiple chambers of the heart. This is a very challenging extension of both the normal 3D-ASM (without update propagation) and SPASM (with update propagation). Several other groups have already been working on complex models of multiple heart chambers, LV-RV [18, 103] or of the complete heart including vascular trunks [32, 104] or of the left half of the heart, including the left ventricle, the left atrium and the aortic outflow tract and the pulmonary veins [105, 106]. It was already mentioned here, that these approaches involve gray level training mechanisms. With the results achieved in this thesis, a next step towards multi-chamber modeling without use of gray level modeling should be taken. In such a multi-chamber model, the first shape variations that will emerge will probably signal relative size differences between the different objects (ventricles, atria and trunks) in the model. If so, such modes may possibly be used for computer aided diagnosis of heart diseases that result in either enlargement or shrinkage of atria or ventricles while other parts of the heart have normal dimensions. Hypoplastic left heart syndrome is such a congenital disease in which the left side of the heart is underdeveloped. Tricuspid atresia and pulmonary atresia are characterized by the absence of the tricuspid valve and pulmonary valve respectively. In both cases, the right ventricle may stay small and underdeveloped. Such shape abnormalities may then be read from components of the b-vector (see Eq. 4.5) of the final model instance after matching.