reconstruction de formes pour la chirurgie assistée par ordinateur basée sur l'enregistrement

Markus Fleute

Laboratoire TIMC - IMAG, Equipe: GMCAO
de l'Université Joseph Fourier - Grenoble I

Shape Reconstruction for Computer Assisted Surgery based on Non-Rigid Registration of Statistical Models with Intra-Operative Point Data and X-ray Images

THESE

présentée par

Markus FLEUTE

pour obtenir le titre de

Docteur de l'Université Joseph Fourier

Spécialité : Modèles et Instruments en Médecine et Biologie

Date de soutenance : 03 Octobre 2001

Composition du Jury :

Président :	Bernard Peroche
Rapporteurs :	Isabelle Bloch Christian Roux
Examinateurs :	Stéphane Lavallée (Directeur de thèse) Laurent Desbat (Codirecteur de thèse) Rémi Julliard

Abstract
This thesis addresses the problem of reconstructing 3D anatomical surfaces based on intra-operatively acquired sparse scattered point data and few calibrated X-ray images. The approach consists in matching the data with a statistical deformable shape model thus incorporating a priori knowledge into the reconstruction process.

Computing such a statistical model requires prior shape analysis in a given population. A new method based on a generic model of the object is used to segment training shapes and to establish point to point correspondence simultaneously in a set of CT images. Scattered point data are then matched with the statistical model using a non rigid 3D/3D registration algorithm. The application of this method for intra and extrapolation of sparse point data is demonstrated within a system for computer assisted reconstruction of the anterior cruciate ligament. To reconstruct a surface from few calibrated X-ray images the statistical shape model is matched to the object contours segmented on the calibrated X-ray images based on a new non rigid 3D/2D registration method. Experiments are performed on a statistical model of lumbar vertebrae for the clinical application of pedicle screw placement.

It is further shown that hybrid matching combining both, 3D/3D and 3D/2D registration, might be an interesting option for certain Computer Assisted Surgery Applications.

Keywords: Computer Assisted Surgery, Medical Imaging, Data Fusion, X-Ray images, Scattered Point Data, Shape Reconstruction, Shape Analysis, Deformable Model, Statistical Shape Model, Model Based Segmentation, Non-Rigid 3D/3D Registration, Non-Rigid 3D/2D Registration

Résumé
L'objectif de cette thèse est la reconstruction de surfaces anatomiques à partir d'un nombre restreint de radiographies et de points acquis en phase per-operatoire. L'approche proposée repose sur une mise en correspondance des données avec un modèle déformable statistique afin d'incorporer de la connaissance à priori sur la forme de l'objet à reconstruire.

L'élaboration d'un tel modèle statistique nécessite l'analyse de forme dans une population donnée. Pour cette analyse un modèle générique de l'objet est utilisé afin d'effectuer simultanément la segmentation des structures et la mise en correspondance de points appariés dans un ensemble d'examens tomodensitométriques. La reconstruction à partir d'un nuage de points est effectuée par une méthode de recalage 3D/3D non rigide. L'application de cette technique d'interpolation et d'extrapolation de données incomplètes est montrée dans un système pour la reconstruction du ligament croisé antérieur. Pour la reconstruction à partir de radiographies une méthode de recalage 3D/2D non rigide est proposée afin de mettre en correspondance le modèle statistique avec les contours de l'objet segmenté dans les radiographies calibrées. Des expérimentations ont été effectuées avec un modèle statistique de vertèbres lombaires, en vue de l'application clinique du vissage pédiculaire.

De plus il est montré que la mise en correspondance hybride combinant le recalage 3D/3D et le recalage 3D/2D pourrait être une option intéressante pour certaines applications dans le domaine des Gestes Médicaux Chirurgicaux Assistés par Ordinateur.

Mots clés: Gestes Médico-Chirurgicaux Assistés par Ordinateur, Imagerie Médicale, Fusion de Données, Radiographies, Nuage de Points, Reconstruction de Formes, Analyse de Forme, Modèle Deformable, Modèle Statistique de Forme, Segmentation Basée sur un Modèle, Recalage 3D/3D non-rigide, Recalage 3D/2D non-rigide

List of publications related to this thesis

I M. Fleute, S. Lavallée. Building a Complete Surface Model from Sparse Data Using Statistical Shape Models: Application to Computer Assisted Knee Surgery. Medical Image Computing and Computer-Assisted Intervention-MICCAI'98. Boston, USA, 1998.

II M. Fleute, S. Lavallée, R. Julliard. Incorporating a statistically based shape model into a system for computer-assisted anterior cruciate ligament surgery. Medical Image Analysis, 1999.

III L. Desbat, M. Fleute, G. Champleboux and O. Desaint. 3D reconstruction using the PIXIUM 4600 digital X-ray detector. International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine Netherlands, 1999.

IV M. Fleute, S. Lavallée. Nonrigid 3D/2D registration of Images using statistical models. MICCAI'99, Cambridge.

V L. Desbat, G. Champleboux, M. Fleute, P. Komarek, C. Mennessier, B. Monteil, T. Rodet, P. Bessou, M. Coulomb and G. Ferretti. 3D interventional imaging with 2D x ray detectors. MICCAI'99, Cambridge.

VI M. Fleute, S. Lavallée, L.Carrat. Nonrigid 3D/2D Registration Based on Statistical Shape Models and its Application to Surgical Treatment of Slipped Capital Femoral Epiphysis (SCFE). CAOS2000, Davos.

VII M. Fleute, S. Lavallée, L.Carrat. Recalage 3D/2D non rigide basé sur un modèle statistique de forme et son application au traitement chirurgical de l'épiphysiolyse de la tête fémorale. Les Journées scientifiques de L'IFR, 9-10 mars 2000, Autrans.

VIII R. Julliard, M. Fleute. Chirurgie du LCA à l'ordinateur: Un plus ou un gadget? Avances et perspectives en matière de gestes medico-chirgurgicaux assistes par ordinateur, Grenoble, 29 septembre 2000.

IX J. Demongeot, M. Fleute, T. Hervé, S. Lavallée. Educational and Research in Medical Optronics in France, ETOP'99 (Education and Technology Transfert in Optronics), SPIE 3831. 2000.

X M. Fleute, L.Desbat, R.Martin, S. Lavallée, M. Defrise, X. Liu, R. Taylor. Statistical model registration for a C-arm CT system. Accepted for IEEE Medical Imaging Conference 2001, San Diego, USA.

XI X. Liu, M. Defrise, L. Desbat, M. Fleute. Cone-beam Reconstruction for a C-arm CT system Accepted for IEEE Medical Imaging Conference 2001, San Diego, USA

Patents

Title: Three-Dimensional Statistical Reconstruction of Surfaces
Inventors: M. Fleute, S. Lavallée, L. Desbat
Patent Cooperation Treaty (PCT) Publication Number: WO 01/22368 A1

Acknowledgements

First I would like to thank my Ph.D. supervisor Stéphane Lavallée. He gave me the opportunity to work in the exciting field of Computer Assisted Surgery and inspired me with his enthusiasm, inspiration and creativity. I would like to express my special thanks to my cosupervisor Laurent Desbat for his time, his patience and his calm and friendship. I would like to thank my thesis committee members, Isabelle Bloch and Christian Roux for their feedback on earlier drafts that made for a much improved final product. I also would like to thank the jury president Bernard Peroche and Rémi Julliard for the fruitful discussions.

For creating a pleasant and motivating working atmosphere I want to thank all TIMC laboratory members especially Jocelyne Troccaz and Philippe Cinquin. I am indebted to my many student colleagues of the Computer Assisted Medical Interventions Department for providing a stimulating and fun environment in which to learn and grow. I am especially grateful to my office mate Christian Huberson who shared with me moments of joy and encouraged me during the occasional bouts of frustration I have experienced during this work. My special gratitude goes also to Sonia Pujol who helped and encouraged me particularly during the last months of this work.

I wish to thank Prof. Merloz and the staff of the orthopaedic department of Grenoble Hospital for the excellent collaboration and fruitful discussions. I am grateful to the staff of the Anatomical department of Grenoble Hospital for providing the specimen of femurs and giving me the opportunity to perform experiments. I am grateful to the Radiological department of Grenoble Hospital for their collaboration.

I also want to thank Prof. Russell Taylor at Johns Hopkins University for giving me the opportunity to work in his research laboratory. I also wish to thank Laurence Vertallier, Gelu Ionescu and Guillaume Champleboux. Parts of this thesis have been inspired by their work.

The Philips research laboratories in Hamburg, the Helmholtz Institut in Aachen and Johns Hopkins Medical School in Baltimore are gratefully acknowledged for providing CT data. Financial support from the the european projects IGOSII, CRIGOS and MI3 is also gratefully acknowledged. Finally I would like to thank the region Rhône Alpes for granting me a scholarship. Without this support this work would not have been realized.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

There are about 15000 Orthopaedic surgeons in the European Community (EC) performing more than 1,000,000 interventions per year. This number is increasing due to several social and economic phenomena such as increasing life expectancy and increasing sportive activity for instance. In spite of the highly trained surgeons there are still important failure rates with large discrepancies of success between surgeons and hospitals. More precisely current limitations of surgical practice occur ([LCT97]).

- Limited ability to implement surgical planning based on pre-operative images or to integrate adequate medical imaging devices into Operating Room (OR) practice.
- Lack of tools, sensors or measurement devices allowing to measure accurately intra-operative data such as the position of bones, tool and cutting guide orientation or location of implants.

In current practice pre-operative image information is transferred to the OR through the surgeon's sometimes limited ability to synthesize and integrate complex *3-D information*. To overcome the limitations and thus to improve the patients' outcomes Computer Assisted Surgery (CAS) systems have begun to emerge from the laboratories and are being used in clinical practice.

The methodology of existing CAS systems follows the classical Perception-Decision-Action paradigm. At the Perception level, multi-modal data are acquired (pre-operative data and models, intra-operative data, post-operative data) and intrinsically calibrated. For imaging devices this means that the raw sensor data are transformed into standardized geometrical entities and expressed in a coordinate system associated with the sensor. At the same stage the image data are segmented (that means labeled and classified) and modeled. At the Decision level, the medical strategy is defined on the basis of all the available models and data. At the Action level, the strategy is carried out through the use of passive, semi-active or active guiding systems. A flow diagram of this methodology is represented in Fig.1.1.

Perception is performed both pre-operatively and intra-operatively. The relevant information mostly originates from medical imaging devices : Computed Tomography

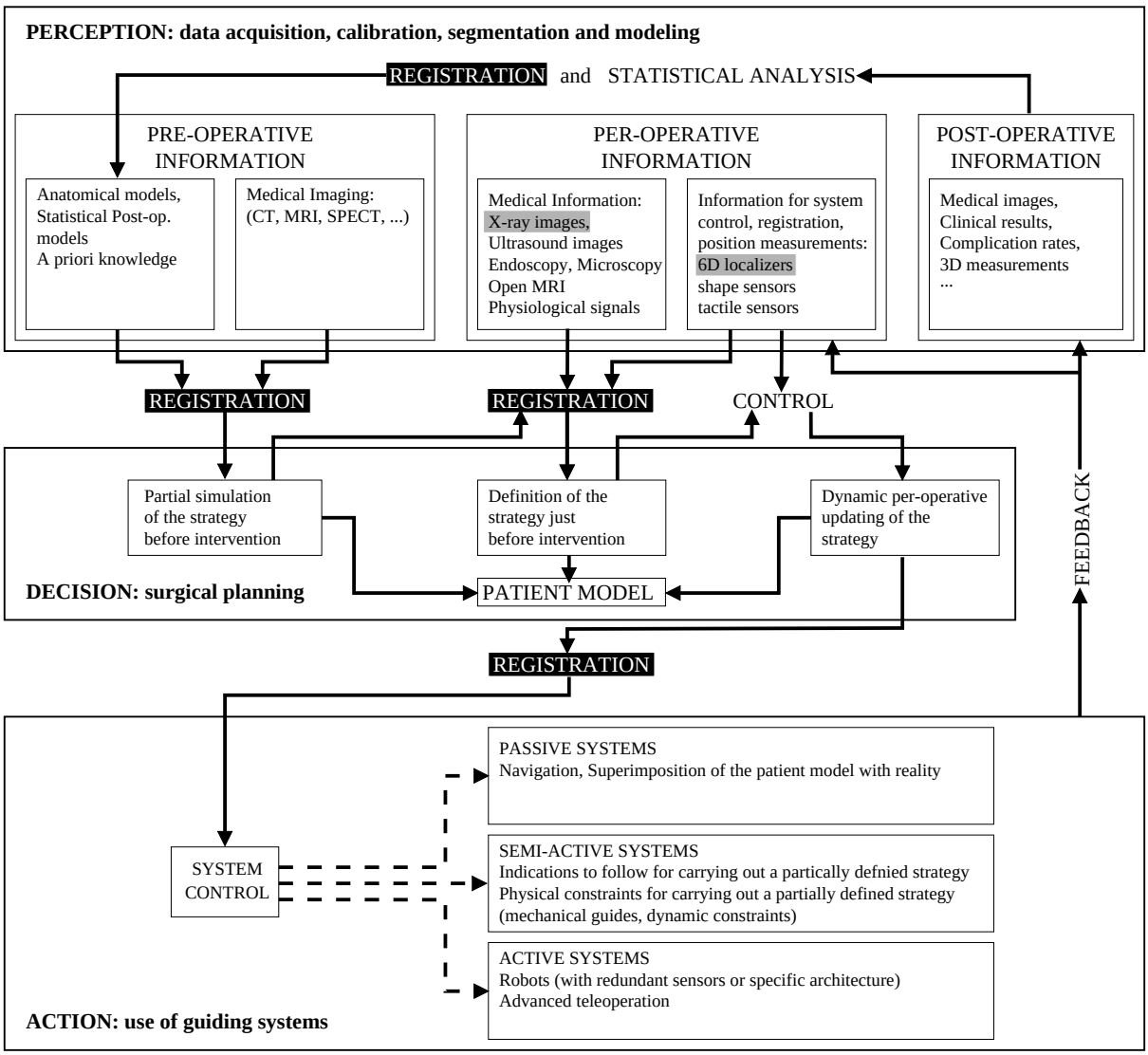


Figure 1.1: CAS methodology follows the Perception / Decision / Action paradigm (adapted from [LCT97]

(CT), Magnetic Resonance Imaging (MRI), Digital Radiography, Ultrasound Imaging (US), endoscopic video images are the most frequent examples. Information originating from other types of sensors such as video cameras, 3D localizers, optical shape sensors, is also available.

The different imaging devices and sensors provide different kinds of data with different dimensionality:

- projections (X-rays), 2D
- sections (ultrasound), 2D
- series of sections or volumes (CT, MRI), 3D
- anatomical surfaces (shape sensors), 3D
- scattered point data (3D localizer), 3D

Breakthroughs in this stage include development of new sensors, taking surgical constraints into account, and accurate calibration of the imaging devices. For instance, the introduction of 3D localizers in the Operating Room (OR) opened the way for many passive navigation techniques.

Decision is an important step involving both construction of patient models and intervention planning based thereon. Construction of patient models means merging all the available information in order to build a virtual patient and implies accurate registration of all information sources, to make optimal use of each one.

Intervention planning is the modeling of an intervention and the simulation of its morphological and functional consequences. Simulation has a *visual* component (navigation through complex anatomical structures) and a *gesture* component (based on modeling the interaction between medical instruments and the human body). The result of this simulation is the selection of an optimal strategy.

Action consists in guiding the selected strategy while it is performed. This implies that the patient model needs to be registered to physical space. Thus when intra-operative sensors are combined with the already created "virtual patient", the surgeon is provided with "augmented reality", which is an effective combined mixture of the real and virtual world. Different levels of assistance exist to perform a selected strategy, including robots in some cases. In all cases, safety is essential, and requires massive redundancy of sensors, processors and actuators. Most existing applications rely on navigation techniques, where surgical tools are tracked, which allows adjustment of their position with respect to the previously defined optimal strategy. In some cases, semi-active systems are used: a robot performs part of the strategy (typically positioning of a guide), and the human operator performs the final part (typically, introduction of a surgical tool in the guide carried by the robot, in Stereotactic Neurosurgery, for instance). Few systems are really active (the robot drives an active surgical tool).

Prof. Russell H. Taylor, Director of the NSF Engineering Research Center for Computer-Integrated Surgical Systems and Technology at the Johns Hopkins University Baltimore, MD, USA predicts a similar evolution for CAS-technology as it has been seen already for Computer Integrated Manufacturing in industry:

> *The impact of Computer-Integrated Surgery (CIS) on medicine in the next 20 years will be as great as that of Computer-Integrated Manufacturing on industrial production over the past 20 years. A novel partnership between human surgeons and machines, made possible by advances in computing and engineering technology, will overcome many of the limitations of traditional surgery. By extending human surgeons' ability to plan and carry out surgical interventions more accurately and less invasively, CIS systems will address a vital national need to greatly reduce costs, improve clinical outcomes, and improve the efficiency of health care delivery. (...)*

Historically, stereotactic neurosurgery has probably been the first surgical discipline to benefit from CAS technology. The fact that some accurate mechanical systems had been in use for a long time (stereotactic frames) and the conjunction of the advent of CT imaging and sufficiently powerful low cost computers may explain that phenomenon. Now, the situation has changed and many physicians and surgeons from very different disciplines have realized the potential benefit of CAS technology.

This thesis focuses on CAS technology applied to orthopaedics, Computer Assisted Orthopaedic Surgery (CAOS). Orthopaedic surgery lends itself well to the application of image guided and robotic technology. Many orthopaedic procedures involve the manipulation or machining of bone, the most rigid structure in the body. The skeletal system can be imaged easily using existing diagnostic techniques, such as X-ray images or CT scans which can be converted to 3D computer models that are used for planning and simulations. Because of the inherent rigidity of bone, the location of the skeletal system during surgery can be correlated dynamically and consistently to a pre-operative computer model. Additionally, unlike soft tissues, bones are able to withstand an applied force from tools such as drills or saws without significant deformation. Thus, a computer model can be built that permits meaningful simulation of the natural motions for bones and joints and their modification during operative procedures.

The potential benefits of integrated CAOS technology are manifold:

- *Improved accuracy of the surgical act:* Improving the accuracy of an intervention strongly influences its success rate. Two different aspects of accuracy must be distinguished:
  - First, it is desirable to be able to define a surgical planning accurately. This means that morphological and functional consequences of an intervention must be taken into account quantitatively. CAS systems have the potential to ensure a more accurate model which is used to predict the results of the surgery.
  - The second aspect of accuracy concerns the implementation of the defined surgical planning. This implies to set up guiding systems that are provided with the exact planning information defined by the therapist, allowing to improve the accuracy of the surgeon's act without obstructing him to take advantage of his dexterity.
- *Minimized invasiveness of surgery:* A major goal is to make surgery less invasive, trying to get close to ambulatory treatment, with limited post-operative consequences.

- *Reduced intra-operative and post-operative complication rates and post-operative pain for the patient:* Improving the accuracy of the intervention and decreasing the access size of surgery, obviously tends to limit the post-operative consequences such as the the risks of infections for the patient.

- *Shortened hospital inpatient time, reduced subsequent outpatient expenses:* Performing small incisions in the skin instead of large open surgery decreases the duration of post-operative stay for the patient at the hospital and may return the patient to the work force earlier.

- *Reduced intervention time:* Although the major objective of CAS technology is to improve the intervention quality, an interesting clinical added value is the reduction of the surgery time causing less contamination risks, reduced anesthesia and less tourniquet duration. From an economical point of view it allows to perform more interventions.

- *Reduced X-ray dose delivered to the patient and to the medical staff:* Time exposure to X-rays should be reduced for the medical staff and for the patient by reducing redundant and unnecessary exposures.

- *Decreased variability of results between different surgeons and decreased stress load:* Relieving the surgeon of the duty to mentally integrate and fuse the medical images should reduce the stress load for the surgeon. The data fusion being performed in a more quantitative objective way by the computer, there should be less intra and especially inter surgeon variability with respect to the surgical result.

- *Global cost of intervention:* The cost issue is a somehow superordinated one since it is directly or indirectly influenced by all of the above mentioned points.

  CAS systems should save money because they improve the success rates of interventions. More precisely the savings in terms of reduced OR costs for the hospital and reduced patient morbidity as a result of the decreased blood loss and decreased operating time may be the most important immediate cost savings as a result of the use of CAS technologies. Near-term cost savings may include decreased rehabilitation time and thus decreased hospital stays and decreased loss of time from work. Longer-term cost savings may include delay or elimination of costly re-operations as well as an associated increase in worker productivity over time.

Not all of these goals are (fully) met by current CAS systems.

When considering the aim of minimizing the invasiveness of surgery, several aspects have to be considered. First minimal invasive surgery can only be performed at a cost: The human operator has no longer any direct visual or manual access to the operated organs. His eyes are relayed by various sensors (x-ray intensifiers, ultrasound probes, endoscopes, ...); his hands are relayed by surgical tools. Not all current CAS systems succeed in helping the surgeon optimally in this difficult task. One also has to keep in mind that current CAS systems quite often are still relying on fiducial markers for registration requiring the patient to undergo an additional surgery for pin implantation. This is an additional risk

of infection for the patient which has to be balanced against the minimized invasiveness of the main surgery. It also increases the pain for the patient and the hospital inpatient time.

Although one of the aims of CAS technology is to reduce the time for each intervention current systems tend actually at best not to reduce the intervention time but usually to increase it. This is very often due to sometimes complicated registration procedures between pre-operative images and physical space or intra-operative data.

Regarding a potential reduction of X-ray dose delivered to the patient for instance, a smaller intra-operative dose could be counterbalanced by the dose delivered by a CT exam that requires small slice spacing for model construction or registration purpose.

Whether the variability of results among surgeons and their stress load is decreased is strongly dependent on the reliability and ergonomics of the system. Complicated registration procedures between pre-operative images and physical space necessitating several trials before being validated by the system are a potential source of frustration for the surgeon for instance.

Concerning the cost issue, only the future can show whether the above mentioned potential savings outweigh the investment required for hospitals or clinics to purchase (or develop), train on, utilize and maintain the CAS systems.

However, patients demand a higher level of service in terms of newer procedures, which are specifically designed to minimize trauma and recovery time. This means that the new technology will be pulled at least partly by customers (patients) into the market and thus can be used as a marketing instrument as has been seen in Germany and some other countries, where clinics that utilize CAS technologies have seen their business jump (at least temporally) soon after adopting such systems. Those clinics that refuse to adopt these technologies run the risk of being viewed as providing lower quality care.

1.2 Objectives / Scope of the thesis

As mentioned above, one key issue in CAS systems is the availability of patient specific 3D models of the anatomical structures on which the surgery is performed.

To obtain these models and to overcome some of the problems occurring when using **pre-operative** images, some authors have proposed to use these rather "heavy" imaging devices directly in the operating room (see for instance [HB80, KHJH88]), thus abandoning the need for any registration for instance. However, these devices are not easily available in standard operating rooms, and imaging systems of radiological departments already have difficulties to meet current demands. The very impressive and recent developments of open magnetic resonance imaging devices that enable surgeons to perform interventions directly inside the MRI device may favour this solution significantly [JB94], but long-term validation and cost reduction is necessary before these devices become widely accepted in standard surgical rooms.

For many applications the computation of a detailed (and precise) 3D attenuation map is not mandatory, i.e. reconstruction of the organ shape is sufficient. Therefore it is desirable to be able to infer 3D-information from intra-operative data only to facilitate

the navigation within the patient and thus allowing to abandon CT data acquisition (pre- or intra-operatively) at least for many standard surgical applications.

This dissertation focuses therefore on the development of new methods and algorithms for the reconstruction of 3D anatomical surfaces for multiple purposes such as surgical planning and visualization relying on **intra-operative** data only, more precisely on incomplete point data and two or three calibrated X-ray projections.

Both, X-ray images and point data are available in standard CAS environments, but in order to infer the 3D shape of the object of interest it is necessary to incorporate a priori knowledge into the reconstruction algorithm. The idea in this work is to formulate the reconstruction problem as a matching problem between a statistical shape model and the intra-operative data.

Two-dimensional statistical shape models have been in use for years now and are well established. Two crucial problems have prevented them so far from being widely used in the 3D case. First, the necessary segmentation of a sufficient high number of training shapes for the statistical analysis, usually available in the form of a CT exam, is a cumbersome and tedious task using today's available manual or semi-automatic segmentation tools. Second, it is necessary to establish point to point correspondence between all training shapes. This is a nontrivial task and becomes manually infeasible. Few known automatic methods address this problem and rely on already segmented images.

The present work proposes a new approach to address both above mentioned problems simultaneously. A generic model of the object is used to segment the training shapes in CT images and to establish point to point correspondence (semi-landmark positioning). A volumetric coarse to fine deformation method based on free form deformations is used to match the generic model to the image data.

In many existing CAS-systems optical localizers or laser scanners are used to acquire scattered point data on patients' bone surfaces in order to register physical space with pre-operative images. A new approach is investigated allowing to inter- and extra-polate intra-operatively acquired point data to obtain a complete surface representation of the actual bone. The idea is to fit a statistical shape model to the available data. Its application is demonstrated for a system for computer assisted reconstruction of the anterior cruciate ligament.

X-ray images are the dominating image modality in the operating room. Thanks to his anatomical knowledge the surgeon is used to mentally fuse 2D images taken from different view points. However for many applications this mental registration is not sufficient to obtain all necessary information about the anatomical situation to properly perform the surgery. Another objective of this thesis is therefore to investigate a method to recover the 3D shape of patient bones or organs intra-operatively using a very limited number (2 or 3) of calibrated X-ray images. The proposed method deforms a previously computed statistical shape model to match the contours segmented on the x-ray images. The fitting of the model to the contours is achieved by using a generalization of a standard registration method (Iterative Closest Point Algorithm) to nonrigid 3D/2D registration. Experiments are performed on a statistical model of lumbar vertebrae suggesting the clinical application of pedicle screw placement.

It is further shown in this work that the combination of both data sources - X-ray

images and 3D point data - allowing to perform hybrid registration with the statistical shape model, might be a very interesting option for certain applications.

Obviously only healthy organs and shape pathologies that can be captured by statistical analysis of a population may be reconstructed with these methods. Fractured organs cannot be modeled. However, various interventions could benefit from such methods. In the case of reconstruction of a torn cruciate ligament for instance, there is no pathologic shape variation of the adjacent bones (tibia, femur). Considering pedicle screw placement for spine instrumentation in the case of a vertebra compression fracture for instance, the shape of the vertebrae the screws are attached to (which are adjacent to the fractured one), is not pathologic; at least there is no pathologic shape variation associated to the reason for the surgery - the fracture. The possible benefits range from reduced radiation dose delivered to the patient, over decreased intervention time and overall inpatient time to the fact that the development of CAS systems with a more favourable cost/benefit ratio would be facilitated. This would help increasing the number of cases where sophisticated CAS technology is applicable *and* affordable.

The development of dedicated hardware involving a light X-ray imaging system similar to a standard C-arm, using a new distortion free flat 2D silicon X-ray detector is part of a European project aiming at developing an innovative integrated system for Minimal Invasive Image Guided Surgery. This project is carried out in the TIMC lab in collaboration with several other European scientific institutions and industrial partners [mi3]. For experiments carried out in the course of this thesis a first prototype of this system is used.

1.3 Chapter Overview

The thesis is further outlined as follows: Chapter 2 presents two clinical orthopaedic scenarios with a significant potential to benefit from the contributions of this thesis in the future. For both of them industrial CAS systems are yet available and some of them will be presented exemplary in order to point out their limitations and the potential added value provided by the proposed methods. Chapter 3 presents the main CAS specific hardware components on which the proposed methods for intra-operative shape reconstruction rely and that are used for experiments carried out within the context of this thesis. This specific hardware comprises an optical localizer system, an interventional X-ray imaging system (image intensifier, C-arm) and a new flat panel X-ray detector. Chapter 4 provides a brief introduction to basic concepts of registration methods, segmentation and deformable models. Special attention is paid to the statistical shape model by Cootes and Taylor whose 3D version is used throughout this work. The following three chapters comprise the main contributions of this dissertation. Chapter 5 presents a new method for highly automated model based extraction of bony structures from CT images. The proposed method is based on the fitting of a generic shape template to the data and aims at establishing simultaneously anatomical point-to-point correspondence between the shape template and the data. Thus statistical 3D shape analysis of large unsegmented image databases becomes more feasible in practice. Chapter 6 presents a method for shape template based surface reconstruction based on scattered point data, and then investigates

a new method for Non-Rigid 3D/3D Registration of intra-operatively acquired scattered point data with a statistically based shape model and its application to computer assisted reconstruction of the anterior cruciate ligament. Chapter 7 investigates how the same statistical shape model can be used for nonrigid 3D/2D registration with contours segmented on few calibrated X-ray views. At the end of the chapter a hybrid approach for surface reconstruction, relying on both point data and X-ray images, is proposed and first experimental results are presented. Chapter 8 provides a summary and gives some indications for future work.

Introduction

1.1 Contexte

Il y a environ 15000 chirurgiens orthopédistes dans la communauté européenne (CE) réalisant plus de 1000000 interventions par an. Ce nombre est en augmentation pour diverses raisons socio-économiques comme l'augmentation de l'espérance de vie et de l'activité sportive par exemple. En dépit du haut niveau de formation des chirurgiens, il subsiste un important taux d'échecs et de grandes variations dans les résultats entre les différents chirurgiens et hôpitaux. Plus précisément la pratique chirurgicale est limitée par :

- des difficultés pour réaliser un planning chirurgical à partir d'images préopératoires ou pour mettre en place des systèmes d'imagerie adéquats dans le bloc opératoire.
- un manque d'outils, de capteurs ou d'appareils de mesure permettant de mesurer avec précision des données peropératoires comme la position des os, des outils et des orientations de guides de coupes ou la localisation des implants.

En pratique courante le chirurgien fusionne mentalement les images préopératoires avec la réalité peropératoire. Toutefois cette fusion mentale est limitée lorsque il s'agit d'informations tridimensionnelles complexes. Afin de surmonter ces limitations et ainsi améliorer les résultats cliniques, des systèmes de chirurgie assistés par ordinateur commencent à émerger des laboratoires et sont utilisés en routine clinique. La méthodologie de ces systèmes suit le paradigme classique de perception - décision - action. Dans la phase de perception des données multimodales sont acquises (données et modèles préopératoires, données postopératoires) et intrinsèquement calibrées. Pour les systèmes d'imagerie médicales ceci implique que les données brutes du capteur soient transformées en paramètres géométriques standards et exprimées dans le référentiel du capteur. Durant cette même phase l'image est segmentée (labellisée et classifiée) et modélisée. Dans la phase de Décision la stratégie opératoire est déterminée à partir de toutes les données et modèles disponibles. Dans la phase d'Action la stratégie est reproduite par l'intermédiaire de systèmes de guidage passifs, semi actifs ou actifs. L'organigramme décrivant cette méthodologie est représenté dans la figure 1.1.

La **perception** se réalise à la fois en phase préopératoire et peropératoire. L'information pertinente provient principalement de systèmes d'imagerie médicales : Tomodensitométrie (TDM), imagerie à résonance magnétique (IRM), radiographies numériques, imagerie à ultrason et images vidéo endoscopiques sont les exemples les plus

fréquents. L'information provenant d'autres types de capteurs comme les cameras vidéos, les localisateurs optiques, les capteurs de forme optiques sont aussi disponibles. Les divers systèmes d'imagerie et de capteurs fournissent différents types de données de dimensions différentes :

- Projections (rayons X), 2D
- Coupes (ultrason), 2D
- Séries de coupes ou volumes (TDM, IRM), 3D
- Surfaces anatomiques (capteurs de forme), 3D
- Nuages de points (localisateur 3D), 3D

Les innovations à ce stade incluent le développement de nouveaux capteurs prenant en compte les contraintes chirurgicales, et le calibrage des systèmes d'imagerie. Par exemple, l'introduction de localisateurs 3D au bloc opératoire a permis l'avènement de beaucoup de techniques de navigation passives.

La **décision** est une étape importante qui comprend à la fois la construction de modèles de patients et le planning de l'intervention. La construction de modèles de patient implique la fusion de toutes les informations disponibles afin de construire un patient virtuel; ceci implique également une mise en correspondance précise de toutes les sources dinformation, afin d'utiliser chacune de ces sources de façon optimale.

Le planning opératoire consiste en une modélisation de l'intervention et la simulation de ces conséquences morphologiques et fonctionnelles. La simulation se décompose en une composante visuelle (navigation à travers des structures anatomiques) et une composante gestuelle (basée sur la modélisation de l'interaction entre les instruments chirurgicaux et le corps humain). Le résultat de ces simulations est la sélection d'une stratégie optimale.

L'action consiste à guider le chirurgien lors de la réalisation de la stratégie optimale sélectionnée. Ceci implique que le modèle du patient soit mis en correspondance avec la réalité opératoire. Ainsi lorsque les capteurs peropératoires sont combinés avec le "patient virtuel", le chirurgien dispose d'une réalité augmentée . Il existe différents niveaux d'assistance pour réaliser la stratégie optimale, y compris dans certains cas l'utilisation des robots. Dans tous les cas, la sécurité est essentielle et ceci implique une redondance importante de capteurs, processeurs et actionneurs. La plupart des applications existantes repose sur des techniques de navigation où les instruments chirurgicaux sont localisés, permettant ainsi d'ajuster leur position par rapport à la trajectoire optimale définie au préalable. Dans certains cas, des systèmes semi-actifs sont utilisés: un robot réalise une partie de la stratégie (typiquement le positionnement d'un guide), et le chirurgien réalise l'autre partie (typiquement l'introduction d'un outil chirurgical dans le guide porté par le robot, en neurochirurgie stéréotaxique par exemple). Peu de systèmes sont réellement actifs (le robot pilote un outil chirurgical).

Prof. Russell H. Taylor, directeur du NSF Engineering Research Center for Computer-Integrated Surgical Systems and Technology à l'Université de Johns Hopkins à Baltimore, Maryland, USA, prédit une évolution similaire pour la technologie de systèmes

GMCAO à celle observée avec l'arrivée du Computer Integrated Manufacturing (CIM) dans l'industrie.

> *The impact of Computer-Integrated Surgery (CIS) on medicine in the next 20 years will be as great as that of Computer-Integrated Manufacturing on industrial production over the past 20 years. A novel partnership between human surgeons and machines, made possible by advances in computing and engineering technology, will overcome many of the limitations of traditional surgery. By extending human surgeons' ability to plan and carry out surgical interventions more accurately and less invasively, CIS systems will address a vital national need to greatly reduce costs, improve clinical outcomes, and improve the efficiency of health care delivery. (...)*

Historiquement, la neurochirurgie stéréotaxique a probablement été la première discipline chirurgicale à bénéficier de la technologie des GMCAO. Le fait que des systèmes mécaniques de précision aient été utilisés depuis longtemps (cadre stéréotaxique), associé à l'arrivée conjointe de l'imagerie TDM et d'ordinateurs suffisamment puissants et peu onéreux, peuvent expliquer ce phénomène. Maintenant, la situation a changé et beaucoup de médecins et chirurgiens provenant de disciplines variées ont réalisé l'intérêt de la technologie des GMCAO.

1.2 Objectifs / Cadre de la thèse

Comme susmentionné, un des aspects fondamentaux des systèmes de GMCAO tient à la disponibilité des modèles tridimensionnels spécifiques aux structures anatomiques du patient, sur lesquels s'exerce la chirurgie. Pour obtenir ces modèles et surmonter certains des problèmes qui se posent lors de l'utilisation des images **préopératoires**, certains auteurs ont proposé d'utiliser ces procédés d'imagerie plutôt lourds directement dans la salle opératoire (voir par exemple , [HB80, KHJH88]), rendant ainsi inutile tout recalage. Toutefois, ces systèmes ne peuvent pas être installés aisément dans toutes les salles opératoires, et les systèmes d'imagerie des départements de radiologie ont déjà du mal à répondre à la demande actuelle. Les impressionants développements récents des systèmes d'imagerie à résonance magnétique ouverts (qui permettent aux chirurgiens de faire des interventions directement à l'intérieur du système IRM) peuvent permettre cette solution, mais une validation sur le long terme ainsi qu'une baisse des coûts est nécessaire avant que ces procédés soient diffusés dans les salles opératoires. Pour de nombreuses applications, l'élaboration d'une carte d'atténuation à la fois détaillée et précise n'est pas obligatoire, la reconstruction de la forme des organes suffit. Il est ainsi souhaitable de pouvoir déduire de l'information 3D uniquement à partir des données peropératoires pour faciliter la navigation à l'intérieur du patient et permettre ainsi d'abandonner l'acquisition de données CT (pré ou peropératoires), au moins pour de nombreuses applications chirurgicales standards.

Cette thèse est dédiée au développement de nouvelles méthodes et algorithmes afin de reconstruir des surfaces anatomiques 3D pour des tâches multiples tels que le planning

chirurgical et la visualisation basée uniquement sur des données **peropératoires**, plus précisément, sur un faible nombre de points palpés et 2 ou 3 radiographies calibrées.

Les images radiographiques et les données palpées sont disponibles dans les environnements GMCAO classiques, mais afin de déduire la forme 3D de l'objet visé il est nécessaire d'incorporer de la connaissance a priori dans l'algorithme de reconstruction. L'objectif de ce travail est de formuler le problème de reconstruction comme un problème de recalage entre un modèle de forme statistique et les données peropératoires.

Les modèles de formes statistiques 2D sont utilisés depuis des années maintenant et sont bien établis. Deux problèmes cruciaux ont freiné jusqu'à présent l'extension de leur utilisation au cas 3D. Premièrement, la nécessaire segmentation d'un nombre suffisant de formes d'apprentissage pour l'analyse statistique, disponible habituellement sous la forme d'un examen TDM, est une tâche laborieuse si on utilise les outils de segmentation manuels ou semi-automatiques disponibles aujourd'hui. Deuxièmement, il est nécessaire d'établir une correspondance point par point entre toutes les formes d'apprentissage. C'est une tâche complexe qui devient infaisable manuellement. Peu des méthodes automatiques connues s'attaquent à ce problème et la plupart repose sur des images déjà segmentées.

Le travail présenté ici propose une nouvelle approche pour résoudre simultanément les deux problèmes mentionnés ci-dessus. Un modèle générique de l'objet est utilisé pour segmenter les formes d'apprentissage dans les images TDM et pour rétablir l'appariement des points. Une méthode de déformation volumétrique multi-échelle basée sur les déformations de formes libres est utilisée pour recaler le modèle générique avec les images.

Dans beaucoup de systèmes de GMCAO existants, des localisateurs optiques ou des scanners laser sont utilisés pour acquérir des nuages de points sur les surfaces osseuses des patients afin de mettre en correspondance les images préopératoires avec la réalité opératoire. Une nouvelle approche est étudiée permettant d'interpoler et d'extrapoler des données palpées en phase peropératoire afin d'obtenir une représentation surfacique complète de l'os. L'idée consiste à adapter un modèle statistique aux données disponibles. Une application de cette méthode est réalisée dans le cadre d'un système de reconstruction du ligament croisé antérieur assisté par ordinateur. Les images radiographiques sont la modalité d'image dominante dans la salle dopération. Grâce à ses connaissances anatomiques le chirurgien est habitué à fusionner mentalement des images 2D d'incidences différentes. Cependant dans beaucoup d'applications cette fusion mentale n'est pas suffisante pour obtenir toute l'information importante sur la situation anatomique afin de réaliser l'intervention correctement. Un autre objectif de cette thèse est donc d'étudier une méthode permettant de reconstituer la forme 3D des os du patient en phase peropératoire en utilisant un nombre limité d'images radiographiques calibrées (2 ou 3). La méthode proposée consiste à déformer un modèle de forme statistique calculé au préalable afin de le recaler avec les contours segmentées dans les radiographies. Le recalage du modèle avec les contours est réalisé en utilisant une généralisation d'une méthode standard pour le recalage (Iterative Closest Point Algorithm, ICP) au recalage 3D/2D non rigide. Des tests ont été réalisés sur un modèle statistique de vertèbres lombaires en vue d'aborder l'application clinique du vissage pédiculaire.

De plus, il est montré dans ce travail que la combinaison de deux sources de données, radiographies et nuages de points 3D, permettant de réaliser un recalage hybride avec le

modèle statistique, pourrait être une option intéressante pour certaines applications.

Bien évidement, seul des organes sains et des formes pathologiques qu'il est possible de modéliser par une analyse statistique d'une population peuvent être construits avec cette méthode. Des os fracturés ne peuvent pas être reconstruits. Cependant, certaines interventions peuvent bénéficier de telles méthodes. Dans le cas de la reconstruction d'un ligament croisé antérieur rompu, par exemple, il n'y a pas de variations pathologiques de forme des os avoisinants (tibia, fémur). En ce qui concerne le vissage pédiculaire dans le cas d'une vertèbre fracturée en compression par exemple, la forme des vertèbres dans lesquelles le vis sont insérées (qui sont adjacents à la vertèbre fracturée), ne sont pas pathologiques. Les avantages de cette méthode sont nombreux. Il vont de la réduction de la dose délivrée au patient, à la diminution du temps d'intervention et de la durée d'hospitalisation jusqu'au développement des systèmes de GMCAO avec un rapport coût/bénéfice plus avantageux. Ceci pourrait aider à accroître le nombre des cas où une technologie de GMCAO sophistiquée serait applicable et envisagable.

Le développement de matériel approprié comprenant un système d'imagerie radiographique mobile, comparable à un amplificateur de brillance classique et utilisant un nouveau détecteur numérique plat 2D sans distorsion, fait parti d'un projet européen dont l'objectif est le développement d'un système intégré novateur pour la chirurgie mini invasif guidée par l'image (projet MI3). Ce projet est conduit par le laboratoire TIMC en collaboration avec plusieurs autres institutions scientifiques et partenaires industriels. Pour les travaux réalisés dans le cadre de cette thèse, un premier prototype de ce système a été utilisé.

1.3 Organisation des chapitres

Les travaux de cette thèse sont présentés de la manière suivante: le chapitre 2 présente deux exemples d'applications cliniques qui peuvent potentiellement profiter des contributions de cette thèse dans un avenir proche. Pour ces deux applications, des systèmes de GMCAO sont commercialisés et certains d'entre eux sont présentés comme exemple, afin de montrer leurs limites et la possible valeur ajoutée apportée par les méthodes proposées. Le chapitre 3 présente les principaux composants spécifiques aux GMCAO sur lesquelles reposent les méthodes proposées pour la reconstruction de formes en phase peropératoire, et qui sont utilisés dans les tests réalisés dans le cadre de cette thèse. Ce matériel spécifique comprend un localisateur optique, un système d'imagerie radiographique interventionnel (amplificateur de brillance) et un nouveau détecteur plat numérique de rayons X . Le chapitre 4 fournit une brève introduction aux concepts de base des méthodes de recalage et de segmentation, et aux modèles déformables. Une attention particulière sera apportée au modèle de forme statistique de Cootes et Taylor dont la version 3D est utilisée tout au long de ce travail. Les 3 chapitres suivants constituent les contributions principales de cette thèse. Le chapitre 5 présente une nouvelle approche d'extraction quasi automatique de structures osseuses à partir des images TDM. Cette méthode est basée sur l'ajustement d'un modèle générique aux images et a pour objectif d'établir simultanément des points anatomiques entre le modèle générique et l'examen scanner. Ainsi l'analyse statistique de formes 3D de base de données d'images non segmentées devient plus facile. Le chapitre 6

présente une méthode pour la reconstruction des surfaces à partir d'un nuage de points. Ensuite une nouvelle méthode est étudiée pour le recalage 3D/3D non rigide entre un nuage de points acquis en phase peropératoire et un modèle statistique de forme, et son application est montrée pour la reconstruction du ligament croisé antérieur assisté par ordinateur. Le chapitre 7 étudie la possibilité d'utiliser le même modèle statistique pour le recalage 3D/2D non rigide avec des contours segmentés sur quelques radiographies calibrées. A la fin de ce chapitre une approche hybride pour la reconstruction des surfaces, reposant à la fois sur des points palpés et des radiographies, est proposée et décrite, les premiers résultats expérimentaux sont aussi présentés. Enfin le chapitre 8 résume les travaux effectués et propose quelques indications pour des développements futurs.

Chapter 2

Orthopaedic Scenarios

2.1 Introduction

In this chapter two exemple clinical orthopaedic scenarios are presented. For both of them industrial CAS systems are yet available and some of them will be presented in order to point out their limitations and the potential added value provided by the contributions of this thesis.

2.2 Reconstruction of the Anterior Cruciate Ligament

The anterior cruciate ligament (ACL), connects the tibia to the femur at the center of the knee. Its function is to limit rotation and forward motion of the tibia with respect to the femur. The rupture of the ACL has become one of the most common injuries among young athletes [SF93]. Because of the instability of an ACL deficient knee and the risk of secondary damage, an increasing number of them needs to have reconstructive surgery. During surgery a substitute ligament is implanted between the femur and the tibia to restore stability. Most frequently, ligamentous tissue of the patient is used, such as the middle third of the patella tendon or the hamstring tendon of the injured knee (Fig 2.1).

In recent literature the accuracy of graft placement has been shown to influence the longevity of ACL reconstruction [ea94]. There is a general consensus that the graft should be placed according to two main criteria [FLS+97, HT93, HB95, ABG+97, MKG95, Mor95, JLD98].

First, the lowest possible anisometry should be obtained, which corresponds to the smallest change in length between the tibial attachment point T and the femoral attachment point F (Fig. 2.2) as the knee moves through the functional range of motion.

With weak anisometry, the graft is subjected to nearly constant tensile forces. Therefore, the risk of rupture during extension or flexion is reduced, and knee stability is improved. In reality, the graft is not of uniform diameter; it approximates the form of a cylinder that widens at its extremities. However, if the centers of the femoral and tibial attachment sites are nearly isometric, then at least the central part of the graft is exposed

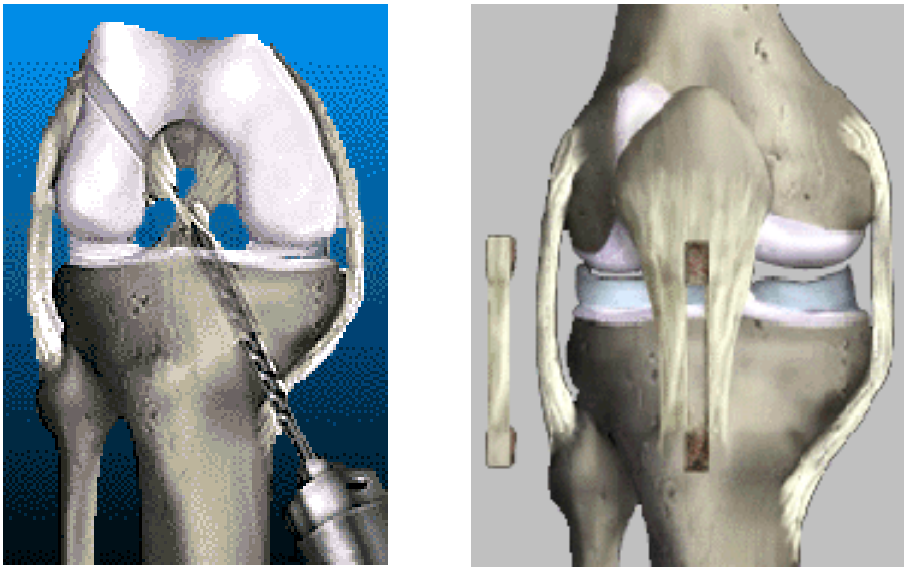


Figure 2.1: ACL reconstruction; drilling the tunnels (left), using the middle third of the patella tendon to replace the original graft (right) [sco]

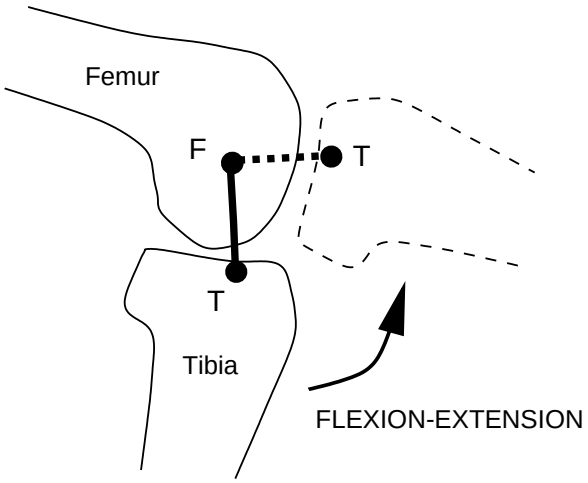


Figure 2.2: Anisometry of the ACL : variation of length between attachment points F and T of the graft along with knee motion

to constant forces.

The second major constraint for the graft placement regards possible notch impingement: If the tibial attachment point lies too far anterior, the graft impinges on the roof of the notch (Fig.2.3). This collision during knee extension may lead to early failure of the implant.

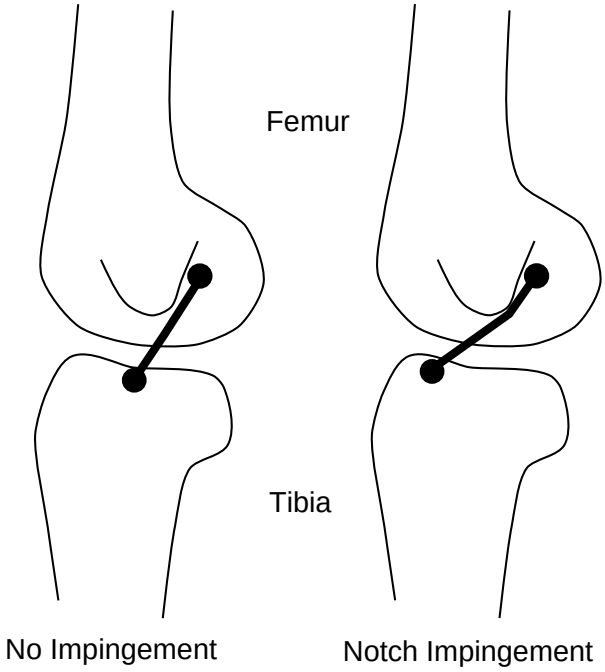


Figure 2.3: Impingement of the graft with the roof of the femoral notch must be prevented

Usually, the surgeon uses mechanical drill guides to achieve reproducible placement of the graft. These instruments rely on freehand technique and visual control, either by open or by arthroscopic viewing, to identify the graft locations. A more objective technique for drilling these holes and placing grafts at desired locations is needed to allow a planning based on quantitative data and thus helping to reduce the variability of this intervention with respect to the clinical results.

Computer Assisted ACl surgery based on pre-operative CT-images and a surgical robot

One commercial system for computer assisted reconstruction of the ACL (developed by Orto Maquet, Rastatt, Germany, now URS Universal Robot Systems, Schwerin, Germany) is based on a pre-operative planning using CT scans of both the injured and the uninjured contralateral knee joint. The planning that means the determination of the position of the graft attachment sides at the tibia and at the femur is based on the relative position of these two bones to each other at the uninjured contralateral side in hyperextension. By this procedure mal-positioning of the graft caused by pathological hyperextension and mal-rotation during the intra-operative planning on the unstable injured side shall be

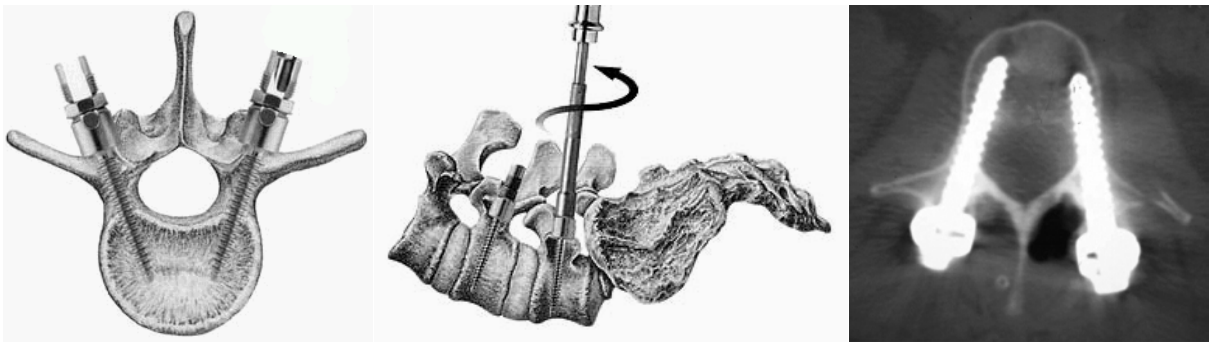
avoided. The CT-scan is acquired after positioning of a metallic registration pin in the femur and tibia in the injured side. The planning itself is accomplished by either using predefined templates or free navigation relying on anatomical landmarks visible in the CT scan. Subsequently the planning of the graft position is transfered to the injured knee joint by manual alignment and carried out by a surgical robot (CASPAR-robot, Computer Assisted Surgery And Planning), see figure 2.4.

Figure 2.4: The CASPAR robot

2.3 Pedicle Screw Placement

Transpedicle screw insertion (see figure 2.5) is commonly used for rigid segmental fixation for various spinal disorders, including scoliosis (abnormal lateral curvature of the spine), and fractures. Unfortunately, the variability in width, height, and spatial orientation of spinal pedicles makes pedicle screw insertion a delicate surgical intervention, especially in the surgical procedure of scoliosis, and for thin pedicles. Surgical techniques are complex relying on image intensifiers and anatomical landmarks. A thorough knowledge of the regional anatomy is mandatory because deformed vertebrae have skewed anatomical landmarks.

Pilot holes are prepared and screws are inserted into the pedicle without any direct visual control. A slight error in direction may result in a significant error in the position of the tip of the screw. X-ray monitoring is possible but very difficult in practice (e.g. control of the screw position with respect to the spinal canal on lateral and AP X-ray images may lead to misjudgment), see figure 2.6. In addition to mechanical instability of the screws, the complications associated with misplaced pedicle screws in the lumbosacral spine are mostly neurological, or vascular [MAS+91]. Previous studies of surgical procedures have



a) b) c)

Figure 2.5: Pedicle screw placement

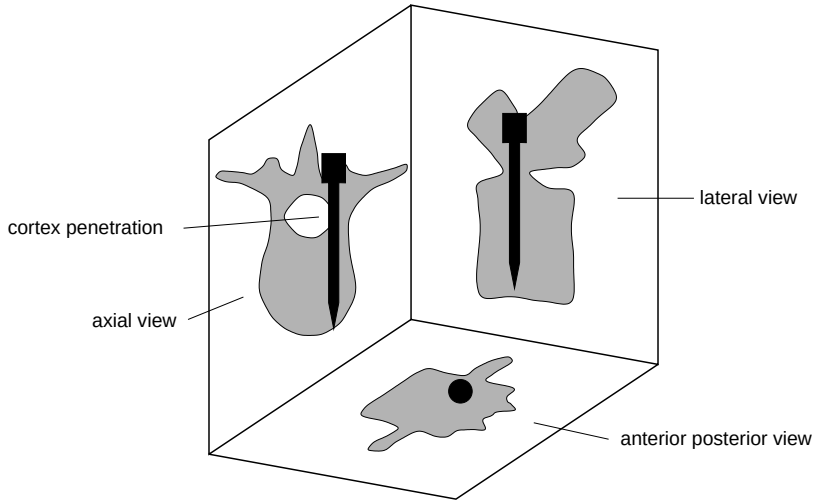


Figure 2.6: Lateral X-ray image may lead to misjudgement of the screw position. Only the axial plane provides the appropriate information.

shown a significant rate of incorrect placement of the screws ranging from 10% to 40% [RCSM86, WSS+88, Sim93, JMC+93, Sai95, VRBa95].

2.3.1 Computer Assisted pedicle screw placement based on pre-operative CT images

To overcome the problems associated with pedicle screw placement based on pure image intensifier images, CT-based navigation systems can be used. These systems help the surgeon to look for the correct spatial position and orientation and the correct length of the pedicle screw as well as to avoid crucial anatomical structures by providing 3D information. At the same time intra-operative irradiation of the surgical staff and the patient is reduced, since the image-intensifier is only used to identify the correct vertebrae but not to guide the screw placement itself.

One example of such a CT-based navigation system is the Stealth-Station(Sofamor-Danek, Memphis, USA). The use of the system is divided into a pre-operative planning stage and an intra-operative execution stage. First, the pre-operative CT images are acquired and transferred to the planning computer. Subsequently the surgeon plans the position and orientation of the screws using either resliced CT images for the principle views (frontal, axial, sagittal) or a 3D presentation of the spine extracted from the CT-data (see figure 2.7).

Intra-operatively, the surgeon first attaches a so called dynamic reference frame to the patient's spine allowing to dynamically track its position using an optical localizer. Subsequently the CT-data and associated planning needs to be registered with the physical space using surface based registration methods based on scattered point acquisitions on the (dorsal) vertebral surface. Afterwards the surgeon can carry out the planning using passive freehand navigation.

While undoubtly improving the success rate in pedicle screw placement ([ea01b]) CT-based navigation systems come along with drawbacks. The necessary pre-operative CT-scan is time consuming, relatively expensive and irradiating for the patient and the necessary registration between CT-scan and physical space is not always straightforward.

2.3.2 Computer Assisted pedicle screw placement based on intra-operative fluoroscopic images

The FluroroNav system (Sofamor-Danek, Memphis, USA) relies only on intra-operative fluoroscopic images, that means no pre-operative imaging is required. It allows real-time navigation in several X-ray projections simultaneously with the fluoroscope turned off and removed from the operating scene. This reduces radiation exposure of both the patient and surgical staff. During the intervention, the surgeon first attaches a dynamic reference frame to the patient's spine as it is the case for the CT-based navigation system. Subsequently a sequence of X-ray images from desired view points is acquired using a calibrated image intensifier, which is also tracked with the localizer. Afterwards the surgeon can navigate with his tracked tools in the 2D images, see Figure 2.8).

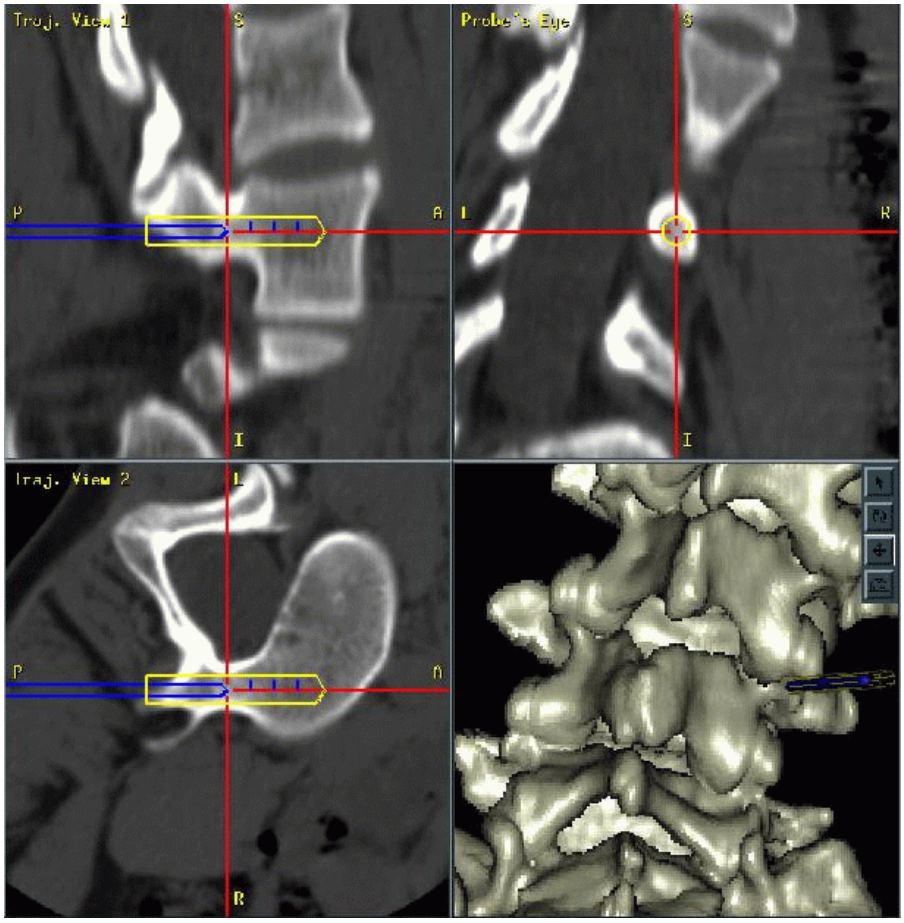


Figure 2.7: The position and orientation of the pedicle screws can be planned preoperatively using a CT scan and a planning computer

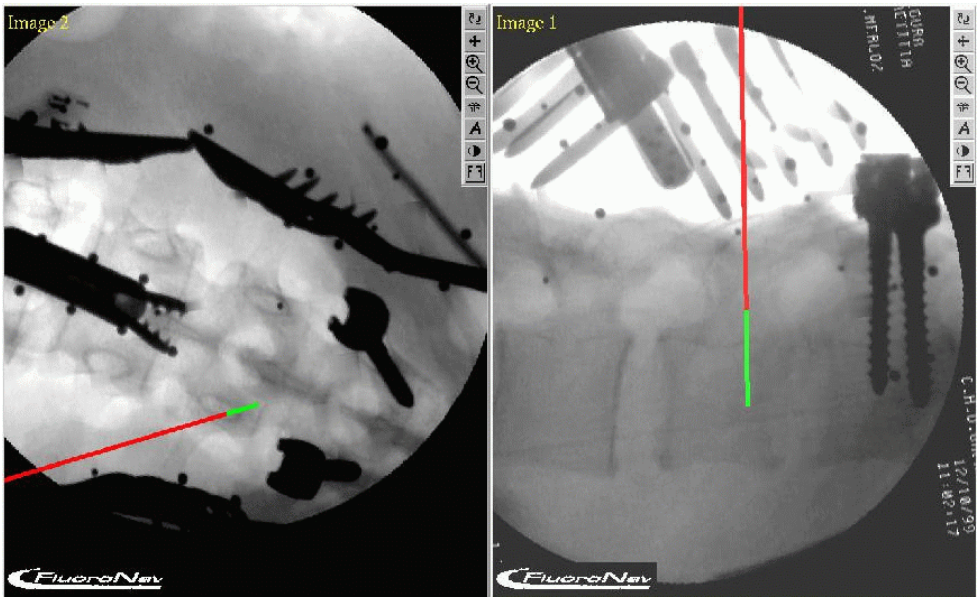


Figure 2.8: The acquired C-arm images can be used for navigation

There is no need for registration between fluoroscopic images and physical space because the transform between both is known during image acquisition by the optical tracker. This results in a gain of time and comfort for both surgeon and patient and decreases potential infection risks. The acquired images reflect the intra-operative anatomy unlike pre-operative CT-images which reflect by nature the pre-operative anatomy. Pre-operative and intra-operative anatomy might differ from each other. On the other hand there is no real 3D information. Specially the axial view which is the most important to watch for hazardous screw placement (see Figure 2.6) is not available.

2.4 Conclusion

Existing CAS systems for knee and spine surgery based on pre-operative images allow precise planning of the intervention. Either passive navigation devices or active robots aid the surgeon to implement the plan. Though providing very precise geometric information the necessary pre-operative imaging is time consuming, expensive and exposes the patient to considerable radiation. Surgical navigation systems based on intra-operative calibrated X-ray images (virtual fluoroscopy) only, aim at overcoming these problems but do not provide often important real 3D information. In many cases knowledge about the 3D shape of the organ is sufficient, i.e. a 3D gray level map as provided by a CT scanner is not required. In this dissertation new approaches are proposed to recover the 3D shape of an organ based on the matching of a deformable shape model relying on intra-operative data only. Chapter 6 presents a method to match a statistical shape model of the femur to intra-operatively acquired point data, applied to a system for ACL reconstruction, which does not require pre-operative imaging. It will be shown that the two main criteria for

correct graft placement, low anisometry and notch impingement can be met relying on intra-operatively acquired cinematic and surface point data extrapolated to a complete model of the bones. In Chapter 7 such a shape model is used to infer the shape of a bone from its partly segmented contours on calibrated intra-operative X-ray projections by matching the model to the computed projection rays. Both methods combine the advantage of pre-operative CT-imaging (real 3D information) with the advantages of intra-operative X-ray imaging (availability in the operating room, low radiation, etc.) for procedures where the expected shape variation of the organ of interest can be captured by a statistical analysis of a population.

Chapter 3

Hardware Components

This chapter presents the main CAS specific hardware components on which the proposed methods for intra-operative shape reconstruction rely and that are used for experiments carried out within the context of this thesis. First a short overview of intra-operative tracking devices is given. Then the basics of interventional X-ray imaging systems are explained and subsequently a new flat panel X-ray detector is described. In the last section the mechanical deformation of the interventional imaging device is discussed.

3.1 Intra-Operative Position/Orientation Tracking Devices

Position tracking devices are fundamental components for virtually all CAS systems and are used during surgery to precisely localize conventional surgical tools, rigid anatomical structures, other medical imaging equipment (e.g., image intensifiers), surgical implants etc. and to capture the kinematics of a limb or to digitize anatomical surfaces. During knee surgery for instance it may be desirable to measure the relative movement of the tibia with respect to the femur. In computer-assisted spine surgery, it might be necessary to determine the position and orientation of a drill relative to a vertebra.

There are a number of attributes which characterize a given intra-operative position sensor including sensing modality, accuracy, resolution, speed, robustness, active vs. passive, and cost. Mechanical sensors determine the position of a sensor endpoint based upon measurements of joint angles and information regarding the kinematics of the device. One advantage of mechanical sensors is their constant activity, i.e they cannot be obscured. On the other hand they tend to be bulky and thus ergonomically challenging for the surgeon. Magnetic sensors measure electrical currents induced in receiver coils when the receiver is moved within a magnetic field generated by the emitter. These sensors suffer from inaccuracies due to other magnetic fields and objects close to the localizer with a magnetic susceptibility $\kappa >> 0$. Acoustic sensors receive signals which are emitted by ultrasonic emitters and determine location via time-of-flight. These sensors are sensitive with respect to temperature shift, humidity and occlusion. The most commonly used tracking devices use optical sensors which track the positions of one or more actively or passively illumi-

nated markers with either three 1D or two 2D cameras(Figure 3.1). They use geometric triangulation to determine the 3D locations of these markers. However optical sensors require a non-occluded line of sight. The achievable accuracy is about $0.1mm - 0.5mm$ (root mean square error between estimated and correct measurement values) [CL98, Sim97]. If a rigid body is equipped with several (at least 3) non-colinearly arranged markers (called dynamic reference base, dynamic reference frame or simply rigid body), it is also possible to determine the spatial orientation of these bodies in addition to their position. These tracking devices are therefore often referred to as 6D localizers. For the experiments conducted in this dissertation two optical localizers are used, $Optotrak^{TM}$ and Active/Passive $Polaris^{TM}$ (Northern Digital Inc. Ontario Canada), see Figure 3.1.

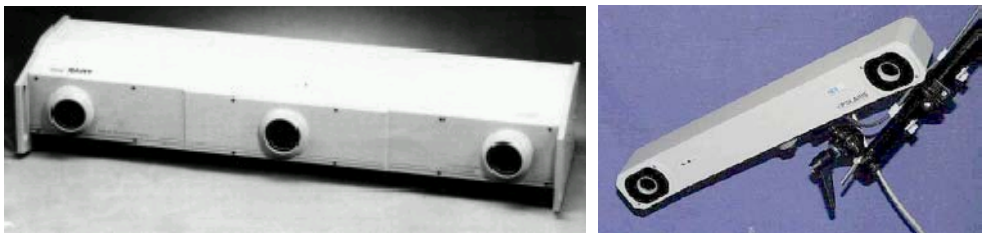


Figure 3.1: Two optical localizers based on 3 1D cameras or 2 2D cameras

3.2 X-ray imaging overview

Conventional medical imaging with X-rays is based on transmission through the patient. X-ray transmission is useful because structures vary in their X-ray absorption, and thus alter the transmitted x-ray beam based on their particular absorption properties. The mono-energetic transmission of an X-ray beam along a path s through an object of thickness t with an (energy-dependent) attenuation coefficient $\mu(x, y, z)$ can be written as:

$$I_d = I_o e^{-\int_0^t \mu(x,y,z)ds},$$

where I_0 is the known intensity of the incident X-ray beam and I_d is the detected intensity. The projected attenuation coefficient is then the natural logarithm of the measured transmission:

$$\int_0^t \mu(x, y, z)ds = \ln \frac{I_0}{I_d}$$

Different technologies exist to measure the transmission. Intra-operatively the most common used technique is based on image intensifiers.

3.3 Interventional X-ray imaging

An image intensifier is a mobile X-ray imaging system comprising an X-ray source and a detector unit, the intensifier. Both are attached to a C-shaped metal arc. The entire

system (C-arm + gantry) usually provides six degrees of freedom, 3 translational and 3 rotational, thus enabling to move the so called fluoroscope in a position according to the spatial location of the object being imaged, see figure 3.2.

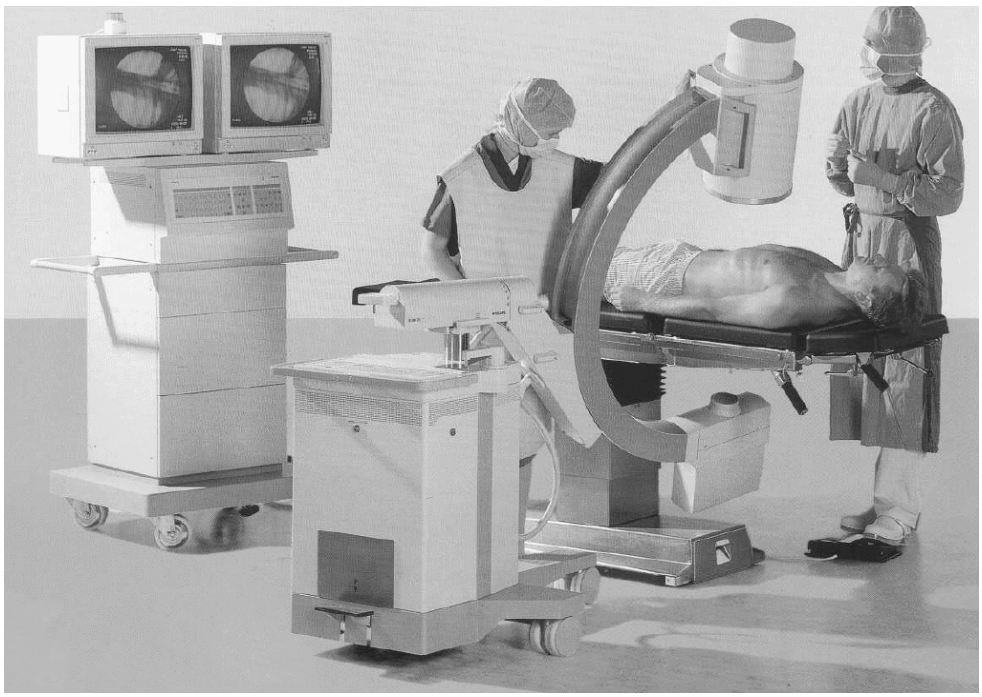


Figure 3.2: Surgical C-arm

Fluoroscopy is routinely used for intra-operative localization of patient anatomy and surgical instrument position.

By providing this information, it facilitates improved accuracy and reduced surgical exposure for a wide variety of procedures when compared to non-image based interventions. Furthermore, fluoroscopy has enabled the development of many interventional and surgical techniques, such as intra-medullary nailing of long bone fractures. The use of fluoroscopy is familiar to most surgeons, particularly in trauma management. It is also very helpful in a variety of spinal procedures. Many spine surgeons routinely use fluoroscopic assistance for the placement of pedicle screws for instance.

3.3.1 Image Intensifiers

Conventional fluoroscopy is performed using image intensifiers that consist of an input fluorescent layer and photocathode, electrostatic lens, anode, and output fluorescent screen. The input fluorescent layer converts the X-rays into light photons, which are then converted to electrons by the photocathode. As the electrons are accelerated towards the anode by a high potential difference, the electrostatic lens focuses them. These electrons are then converted back to photons by the smaller output screen. The combination of geometrical reduction and electron acceleration results in a large gain in brightness for

the image intensifier, which allows the device to be used in the lower dose conditions of fluoroscopy.

Imaging intensifiers, however, also have drawbacks. The output images exhibit a large amount of distortion, due to both the curvature of the input fluorescent layer and the deflection of the electrons by the earth's magnetic field ([ea97], Figure 3.3). Moreover, the latter effect is dependent on the position of the image intensifier, which changes in a system undergoing motion. As in film screen systems a digitization step is required for post-processing techniques. For systems requiring detector motion, the large size and weight of an image intensifier may complicate mechanical engineering requirements. In order to use X-ray image intensifiers for surgical planning or for registration, it is necessary to correct distortions of the images acquired with these sensors.

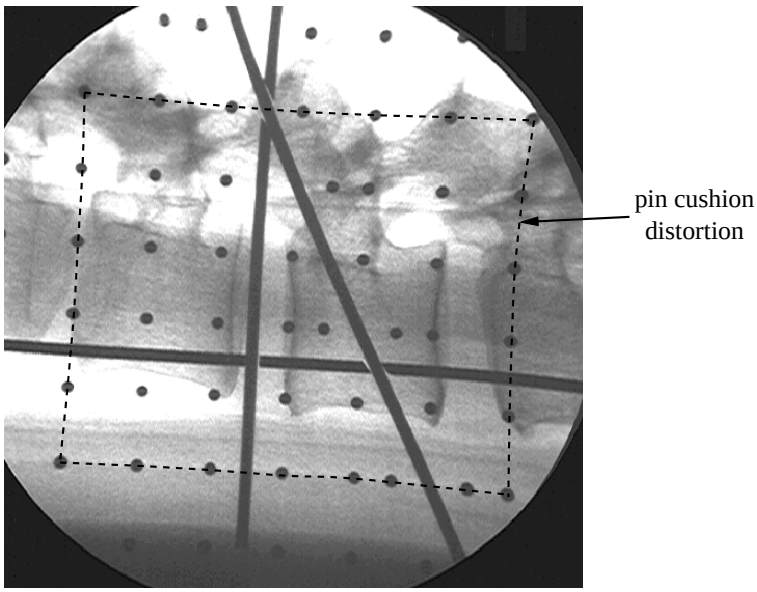


Figure 3.3: Imaged vertebrae with superimposed calibration cage. The pin cushion distortion is visible

3.3.2 Flat panel digital X-ray detectors

Technology issues

To replace film-screen systems and image intensifiers, medical imaging companies have been developing digital distortion-free, flat-panel X-ray detectors, which can be divided into two main groups: direct detectors, and indirect detectors. Direct detectors directly convert X-rays to electrons for signal detection, while indirect detectors include an intermediate step which converts X-rays to light photons (which are then converted to electrons). The electric image is stored in each pixel until the panel is read out, when a transistor acts as a switch to send the pixel charge to an analog-to-digital converter (ADC), and from there to a computer.

The experimental set up used in this dissertation comprises a new 43cm x 43cm digital flat panel detector developed by Trixell, Moirans, France which is dedicated to General Radiography, see Figure 3.4.

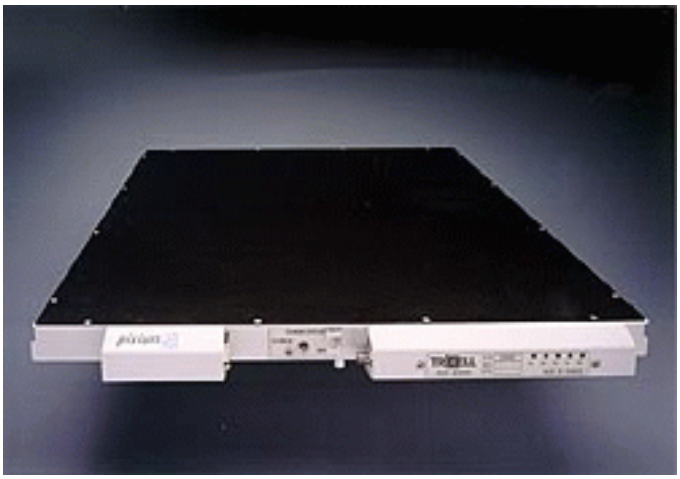


Figure 3.4: The Pixium 4600 digital X-ray detector from Trixell

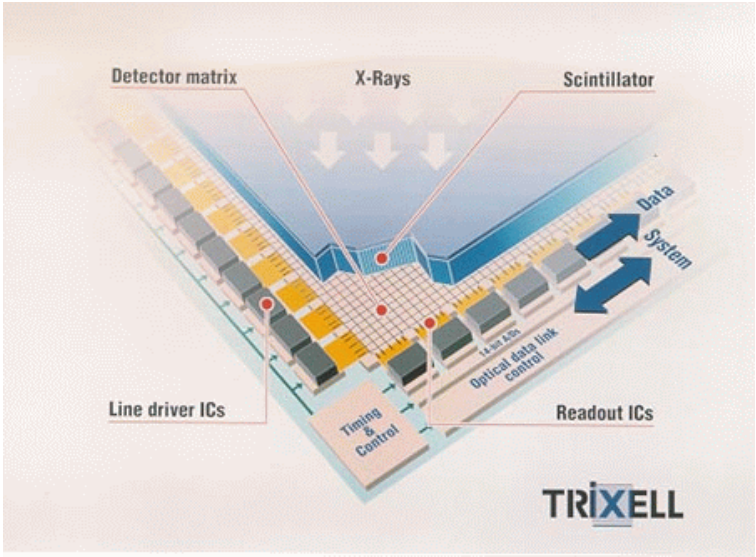


Figure 3.5: Technology of the Pixium 4600 X-ray detector

Figure 3.5 shows the schematics of the (indirect) detector structure. The heart of the flat panel detector is a Cesium Iodide (CSI) scintillating screen, coupled to a 3120 x 3120 pixel, 143μ pitch array of amorphous Silicon (a-Si) diodes deposited on a glass substrate. Sensitive to visible light, these photodiodes are activated by the scintillator covering the matrix, which converts the X-ray quanta. The detective panel is connected to dedicated low noise electronics which provide line addressing and column multiplexing into a serial

electrical signal. The signal is then digitized over 14 bits to provide a direct digital image output, available for the radiographic system via an optical fiber.

Low Level Image Processing

Imperfections in the detector and electronics result in serious imaging artifacts in uncorrected images. Typical imaging artifacts include bad pixels, bad lines, non-uniform readout gain, dark current, and regions of non-uniform X-ray conversion. Thus pre-processing of the images is necessary and includes several basic corrections and associated calibrations. Basically offset correction, gain correction and defect interpolation is applied thus all fixed pattern noise is removed by this pre-processing.

The gain and the dark current (offset) are measured by the collection of two additional images: a flat field , i.e. no object present, and a dark field, i.e. X-ray beam blocked.

At first a pixelwise offset subtraction is performed, subsequently sensitivity normalization is, usually referred to as gain correction is carried out. Each pixel is multiplied with its individual gain factor. The gain correction is based on a linear signal model. Alternatively, a non-linear correction scheme could be used (a second-order model, based on a dark measurement and several flat fields at different intensities for instance).

The last procedure is the defect interpolation. Each pixel that shows unusual behavior is marked in a defect map and is replaced in the final images by the interpolation of its neighbor values in either direction. Unusual behavior in first instance means completely insensitive pixels, but also too large or too low offset, more than certain sensitivity deviation from the average sensitivity (25%).

3.4 Calibration of the imaging system

Before the imaging devices can be used for quantitative measurements and subsequent surgical planning, they need to be intrinsically and extrinsically calibrated.

3.4.1 Intrinsic Calibration

Intrinsic calibration models the imaging system (detector + source) itself, thus providing a function that transforms raw sensor data into geometrical entities, which are expressed in a coordinate system associated with the sensor.

As said before, conventional interventional X-ray imaging systems using an image intensifier suffer from geometric distortions. The correction of these distortions is thus usually part of the intrinsic calibration. Since a fully digital distortion free detector is used for experiments in this work, the intrinsic calibration only implies the determination of the source position in the detector reference system.

The basis for the projection from 3D space into the 2D image is is modeled as a linear cone beam projection model (pinhole camera model), see Figure 3.6. It is assumed that the origin of the source-centered coordinate system coincides with the focal spot of the X-ray source, and the z axis coincides with the system's optical axis, as illustrated in Figure 3.7. f is the effective focal length, i.e. the distance between the X-ray source and

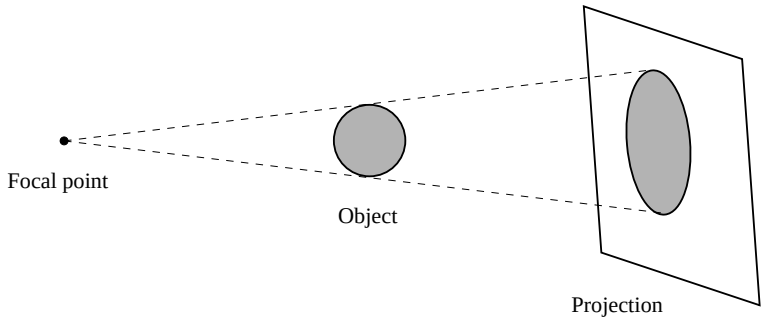


Figure 3.6: 3D cone beam projection model

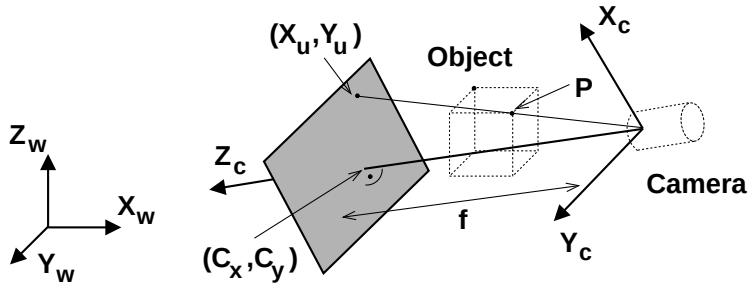


Figure 3.7: Geometry of the perspective-projection camera model

the image plane. The relationship between the position of a point P in world coordinates, and its projection in the detector coordinate system is defined by a sequence of coordinate transformations. The first is a rigid body transformation from the world coordinate system (X_w, Y_w, Z_w) to the X-ray source centered coordinate system(X_c, Y_c, Z_c). This is expressed as

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = R \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix}$$

where R is a rotation matrix and T a translation vector. The second transformation is a perspective projection of point P in the source coordinates to the position of its detector coordinates. This is described by

$$X_u = f\frac{X_c}{Z_c}$$

and

$$Y_u = f\frac{Y_c}{Z_c}$$

where f is the focal length of the pinhole camera.

The projection parameters i.e. the coordinates (x,y,z) of the source position are determined in a calibration procedure. For this purpose a calibration probe is used, which contains 200 steel spheres placed coplanar in a regular grid together with 3 intersecting cylinders (Figure 3.8). This probe is imaged twice by the C-arm for two different precisely known distances to the image intensifier (figure 3.9). Since the calibration probe is tracked by the position sensor, the 3D position of the center of every sphere in the global coordinates system can be calculated. This allows subsequently to calculate the 2D positions of the projections of the spheres in the detector coordinates system. First the exact projection of the 3 cylinders are determined using a Hough transformation (figure 3.10). Then the known geometry of the calibration plate allows in conjunction with dedicated image processing algorithms to determine the sphere center positions in the two images and for each position its corresponding sphere in the calibration plate. Searching for the parameters describing the projection of the spheres into the image plane for both calibration plate positions results in solving an over-determined system of equations. This is accomplished using an iterative error minimizing algorithm; the method is described in detail in [Tsa86].

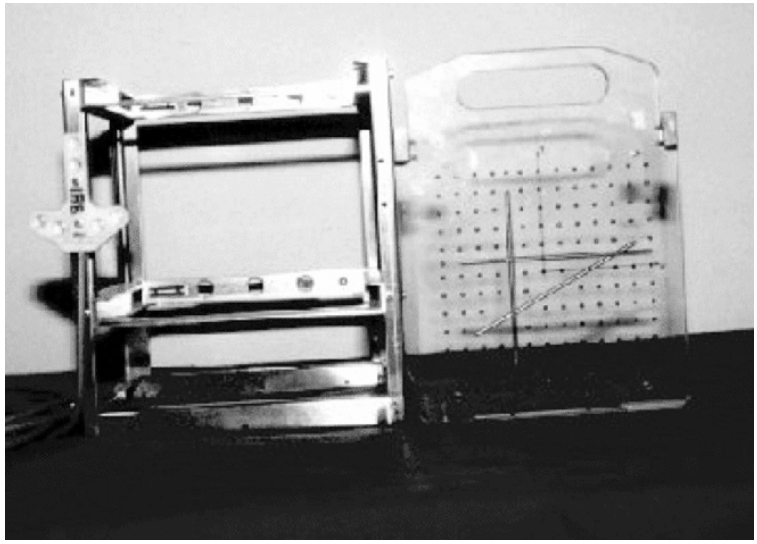


Figure 3.8: The calibration cage

Mechanical deformations of a C-arm

During the C-arm motion the intrinsic geometry of the X-ray imaging system will not remain constant due to the mechanical non-rigidity of the C-arm structure and the weight inequity due to the detector end of the C-arm being heavier than the source. The resulting minor orientation dependent bending of the C-arm affects the optical axis of the X-ray system, which is defined as the normal dropping from the X-ray source onto the detector plane. That means in principle that for each new C-arm position a new intrinsic calibration step is necessary. In fact several commercially available systems do this. Intensifier based

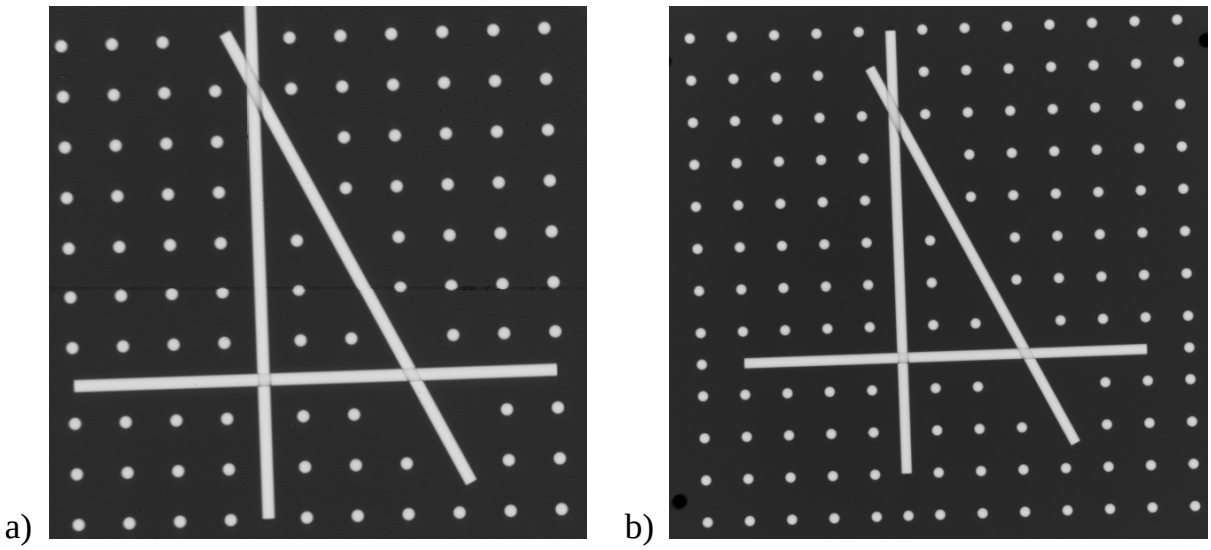


Figure 3.9: Corrected calibration cage images for two different positions

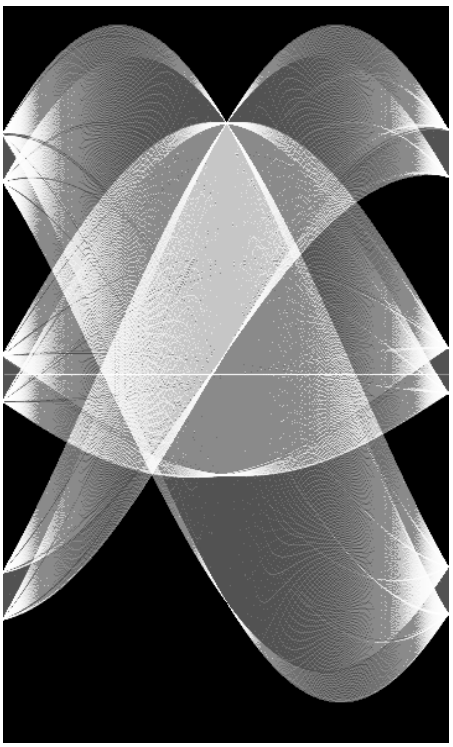


Figure 3.10: Hough transformation of the calibration cage image

C-arm systems for CAS applications do need this calibration step anyway because of the geometric distortions due to the Earth's magnetic field mentioned in section 3.4.1 and the effect of external electromagnetic fields in the vicinity of the C-arm generated by other electrical devices. Therefore CAS systems based on image intensifiers usually have a calibration cage attached to the image intensifier that remains during image acquisition. This further reduces the already limited workspace and also decreases the image quality due to the calibration landmarks superimposed to the object (patient) being imaged.

Since digital X-ray detectors do not suffer from geometric distortions there is no need for such a calibration target. However to account for the mechanical instability, it is necessary to precisely quantify the deformation for each possible position of the C-arm. Assuming that the mechanical deformations are perfectly elastic, the deformation can be measured once offline and subsequently modeled and corrected for during intra-operative image acquisition. Figure 3.11 shows the result of a preliminary study; the source to detector distance as a function of the orbital angle of a prototype of a C-arm based X-ray imaging system equipped with angle encoders and the above presented digital X-ray detector (Figure 3.12).

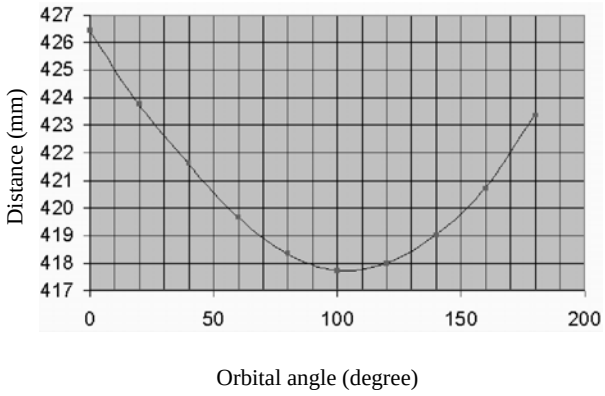


Figure 3.11: Source to detector distance as a function of the orbital angle

3.4.2 Extrinsic Calibration

The extrinsic calibration of the C-arm i.e. the determination of the transformation between the detector coordinate system and the global system wide used reference system is carried out using the angle encoders and an optical tracking system.

3.5 Conclusion

Position tracking devices are fundamental components for CAS systems and most commonly based on optical sensors, either using actively or passively illuminated markers. They are used to track the position and orientation of patient organs, surgical tools, implants etc. Within the context of this thesis the function of such a device is to acquire

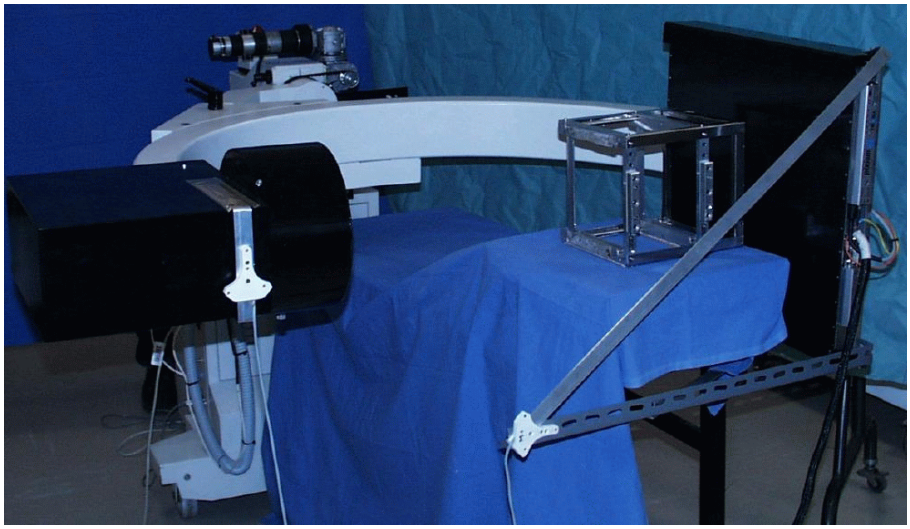


Figure 3.12: First prototype of a new interventional X-ray imaging system dedicated to CAS surgery, equipped with angle decoders and optical markers for position/orientation tracking.

intra-operatively points on a bony surface and to track the position and orientation of an interventional X-ray imaging device (C-arm) equipped with a new digital distortion free flat panel detector. Calibration of the imaging device taking the mechanical deformation of the C-arm into account is necessary for quantitative measurements which is essential for the shape matching method proposed in chapter 7.

Chapter 4

Segmentation, Registration and Deformable models for Image Analysis and CAS

The goal of any image analysis system is to infer information required for a specific task. The inter- and intra-subject variability inherent to biological structures makes medical image interpretation a difficult task in general.

A commonly applied subdivision scheme of image analysis tasks into five abstraction levels according to the degree of a priori knowledge involved in operations at that level is shown in Figure 4.1. The lowest level includes filtering and normalization of the raw images. At the next higher level (low level) segmentation on the filtered image is performed. At the low level vision level geometric scene elements like surfaces, volumes or contours are extracted. At the object recognition level, single objects are recognized and labeled. At the high-level vision level interpretation of entire scenes is performed.

Segmentation is an essential part of any image analysis system. Within the context of this dissertation it is used to extract shapes in 3D gray level images in order to perform subsequent shape analysis (chapter 5) and to define the outline of an object's projection in X-ray images (chapter 7). In section 4.1 some of the most common low level segmentation algorithms are presented and it is discussed why they often lead to unsatisfactory results when applied alone. *Registration* is part of virtually any CAS system (see chapter 1). A brief introduction to this topic is given in section 4.2 following [Lav96, MV98]. Prior to any registration, so-called reference structures have to be segmented in general. Registration is dual to segmentation in such a way that conversely the knowledge of the correct registration (transformation) would facilitate the often difficult segmentation. This issue is discussed further in [HLC98] for instance. *Deformable models* can provide efficient means to take into account this duality and to incorporate a priori knowledge into the segmentation/registration process. This is discussed in section 4.3, where special attention is paid to the statistical shape model by Cootes and Taylor [TCCG92], that provides an attractive framework for image interpretation tasks based on the statistical analysis of object shapes. This model is applied in subsequent chapters for surface reconstruction using sparse data.

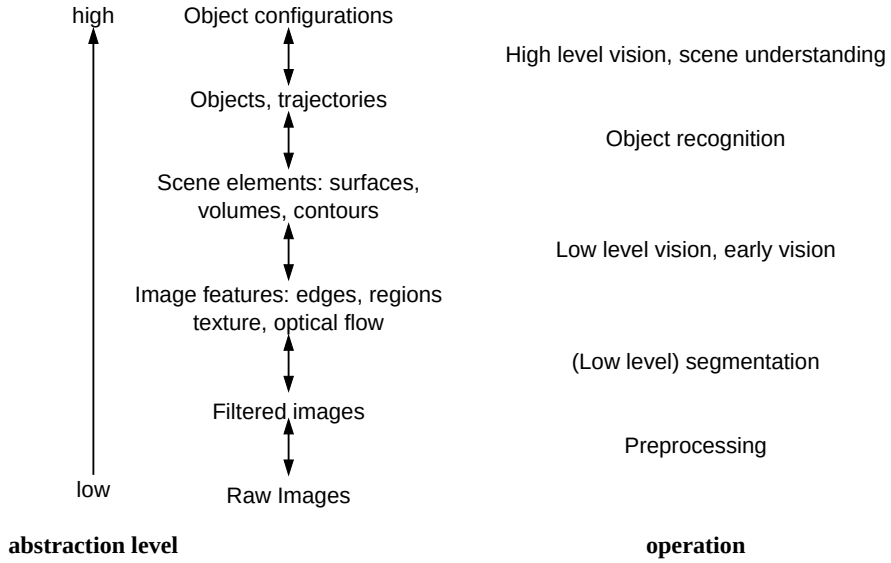


Figure 4.1: General classification of image analysis methods

The last section of this chapter introduces the concept of iterative corresponding point matching which is likewise applied in the subsequent chapters.

4.1 Segmentation

The segmentation of anatomical structures - the partitioning of the original set of image points into subsets corresponding to features characteristic for the structures - is an essential stage of most medical image analysis tasks, such as registration, shape analysis, object recognition, etc.

4.1.1 Manual Segmentation

Most segmentation of 3D medical images is currently performed using manual slice editing, which is extremely labor-intensive and time-consuming. Usually a skilled operator, using a computer mouse or trackball, manually traces the region of interest on each slice of an image volume. Manual slice editing suffers from several drawbacks, such as the difficulty in achieving reproducible results, operator bias, forcing the operator to view each 2D slice separately to deduce and measure the shape of 3D structures, and operator fatigue. Therefore more automatic methods are of high interest. Exemplary some of the more common low-level segmentation methods are summarized in the following section (see also [ACC+96] for further details).

4.1.2 Low-Level Segmentation methods and their limitations

Usually the features to be searched for are either edges or regions. Therefore two conceptually different approaches can be distinguished. Region segmentation consists in searching for homogeneous areas in the image. Edge detection consists in searching local discontinuities of the gray-level function of the image. Region based segmentation methods always yield closed (region) boundaries, which is not the case for edge-finding techniques.

Region based segmentation

Thresholding The most basic segmentation procedure that may be carried out is thresholding of a gray level image. Thresholds may be applied globally across the image (static threshold) or may be applied locally so that the threshold varies dynamically across the image. The success of this approach depends on whether suitable thresholds exist and whether they can be inferred from the image.

Region Growing and Split-and-Merge Algorithms Region growing algorithms start with a number of seed pixels or seed regions and grow these regions by adding to a region previously unassigned neighboring pixels that correspond to some similarity criterion for that region [GW92]. This method proceeds to successively divide an image into smaller non-overlapping regions if some similarity criterion is not met, otherwise no split of that region is carried out. The final result of the splitting is an over-segmented image. A merging procedure is then applied to merge neighboring regions under the same homogeneity predicate that was used for splitting.

Watershed Segmentation For this algorithm the gradient magnitude at each pixel of a 2D image is considered as the height of a surface in 3D, regions are formed by simulating a flooding of the image that begins at local minima of the gradient image function [BM93]. The boundaries of each region stop advancing when neighboring flooding regions meet and the boundaries thus correspond to ridge-paths of the gradient magnitude. The regions defined by the closed boundaries represent an over-segmentation of the image, since there will usually be many such regions corresponding to each object. If the gradients are computed at successively higher scales, the number of local minima (flood basins) in the gradient magnitude image will decrease.

Edge based segmentation

Gradient Operators, Edge and Zero-Crossing Detectors The boundaries of structures in images correspond to local gradient extrema or Laplacian zero crossings of the gray-level function. Due to sensor imperfections, discretization, object irregularities etc. this implies differentiating a noisy signal. Usually the design of an edge detection operator is preceded by defining a set of performance criteria such as one response to one edge, good edge localization, high probability of detecting true edge points and low probability of falsely detecting non-edge points. Canny proposed in [Can86] a derivative of a Gaussian filter as a near-optimal filter with respect to the above criteria. In [Der87]

Deriche proposes a recursive and separable filter with an impulse response similar to that of the derivative of a Gaussian, allowing an efficient implementation especially in the case of large filter masks. Shen and Castan [JC92] propose a derivative filter optimizing the trade-off between edge detection and edge localization.

Thresholding of the filtered image is usually performed in two steps. First, local maxima in the gradient direction are extracted (Figure 4.2 left). Second, hysteresis thresholding is applied in order to eliminate false isolated edge points but keep high connectivity of edge chains (Figure 4.2 right).

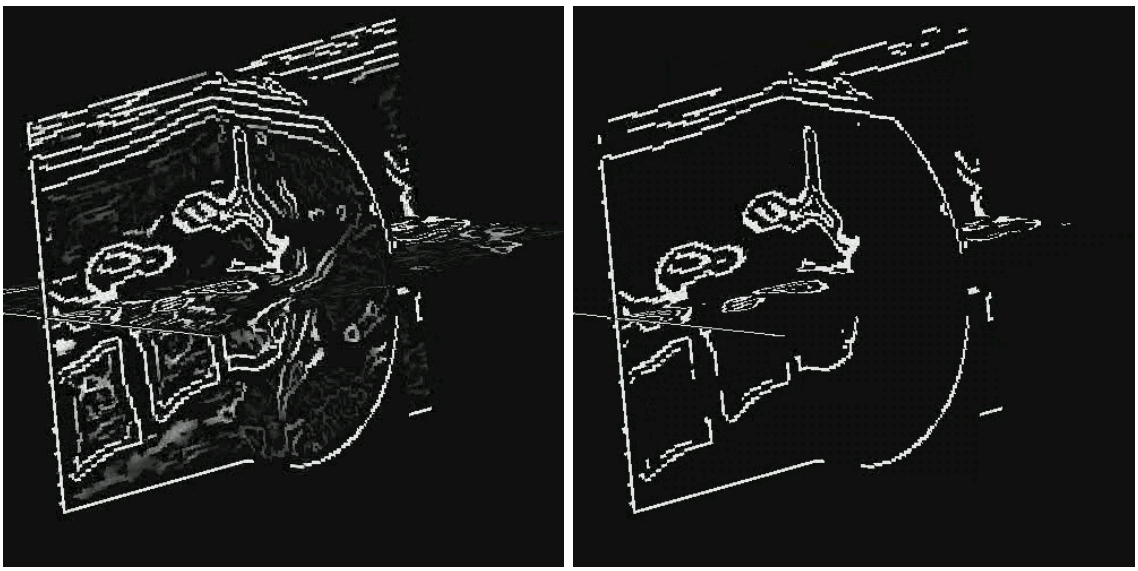


Figure 4.2: After applying a Canny-Deriche filter to an image, thresholding is performed in two steps: First local maxima in the gradient direction are extracted (left). Then hysteresis thresholding is applied in order to eliminate false isolated edge points but keep high connectivity of edge chains (right). Due to considerable variability within the subject perfect edge detection is not possible.

Segmentation using low-level image processing techniques, usually still requires considerable amount of expert interactive guidance or manual post-processing to obtain acceptable results. Furthermore, automating these approaches is difficult because of the shape complexity and variablity within and across individuals and varying imaging conditions. In general, the underconstrained nature of the segmentation problem limits the efficiency of approaches, that consider local information only. Noise and other image artifacts can cause incorrect regions or boundary discontinuities in objects recovered by these methods. Without a priori knowledge medical image segmentation is an ill posed problem and therefore only in specific cases one of the above mentioned methods is able to solve a given segmentation task alone.

4.1.3 Incorporating a priori knowledge

Knowledge can be incorporated into image analysis systems in general in different ways. Procedural (algorithmic) knowledge is incorporated *implicitly* in the algorithm or method to interprete the image. Procedural knowledge representation is not very modular, and little reusable but at the same time very efficient. Declarative (expert) knowledge is represented *explicitly* and symbolically and is independent of the methods that perform inferences on it. It can be represented using formal methods like predicate logic, production rules or semantic nets. Declarative knowledge can be represented in contrast to procedural knowledge as highly modular and is often reusable but on the other hand less efficient. It is also possible to combine both kinds of knowledge representations.

Often methods working at the abstraction level of low level image segmentation are combined with methods working at one of the three highest abstraction levels to exploit (bottom-up) constraints derived from the image data together with (top-down) a priori knowledge about the structure to be analyzed

4.2 Registration

In computer integrated surgery and therapy, registration of all the information available for a given patient i.e. the establishment of a graph of relations between all the coordinate systems involved, is an essential step (see also Figure 1.1) and links together all available information. Figure 4.3 shows a typical CAS scenario where various coordinates systems have to be related to each other:

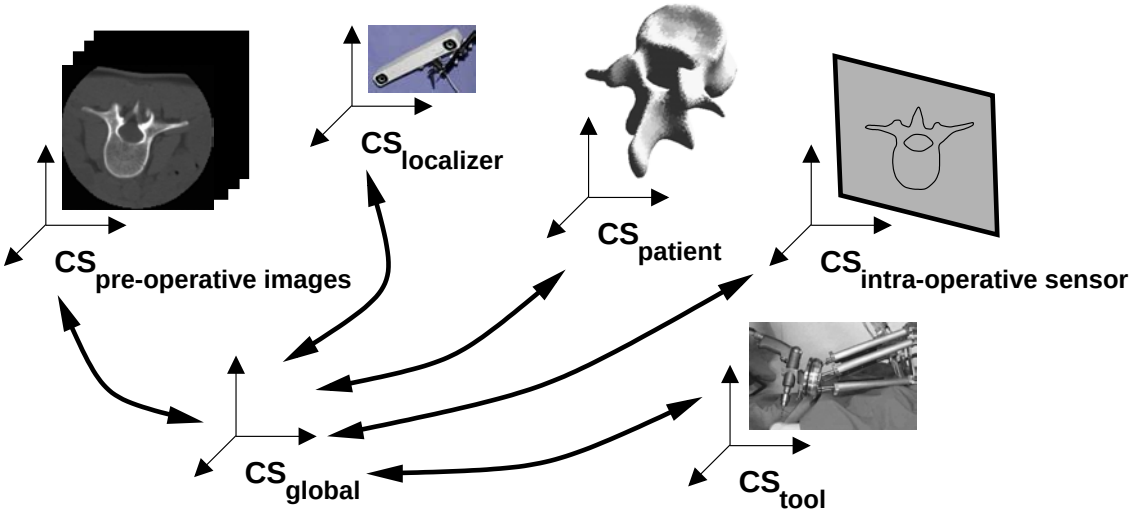


Figure 4.3: Different coordinate systems in a typical CAS setting

- **Medical images or Models/Atlases** with different modalities such as Computed Tomography, Magnet Resonance Imaging, Ultra Sound, X-ray images and different dimensionalities (2D, 3D, 2D + time, 3D + time).

- **Positioning information** gathered by
  - 6D localizers (optical tracking devices), providing the position/orientation of organs or tracked tools (passive guiding systems)
  - position/angle encoders of intra-operative imaging devices, semi-active or active guiding systems (robots).

When considering registration of medical images, one can distinguish between three main application areas:

Intra-operative registration usually aims at estimating an accurate relation between pre-operative images/models and intra-operative data of the same patient thus performing **intra-patient registration**. Differences between data can be due to

- Different viewpoints
- Different imaging modalities
- Different acquisition times

Inter-patient registration refers to registration of data of different patients.

Atlas-patient registration refers to registration between an atlas or a deformable model and data of a patient. It is a generalization of inter-patient registration, if the model/atlas incorporates the same kind of information as the patient data. The problem of registering sparse intra-operative patient data with a statistical shape model can be thought of as being such an application.

According to Lavallee [LCT97], the registration process can be divided into 3 steps:

1. Definition of a relation between coordinate systems,
2. Definition of reference features and definition of a disparity (or similarity) function between them,
3. Optimization of the disparity (or similarity) function.

4.2.1 Definition of a relation between coordinate systems

Definition of coordinate systems

In order to register two modalities of information A and B, the first step is to associate a coordinate or reference system to each of both modalities. The objective is then to estimate the transformation **T** between Ref_A and Ref_B.

Definition of a relation

Limiting to static cases the transformation **T** is a function that transforms coordinates $M_A = (X_A, Y_A, Z_A)$ in Ref_A into coordinates $M_B = (X_B, Y_B, Z_B)$ in Ref_B:

$$M_A = \mathbf{T}(M_B)$$

Nature and domain of the transformation:
The transformation **T** can vary in its nature and regarding its domain.

Rigid transformations: An image coordinate transformation is called rigid, when only translations and rotations are allowed. Rigid-body transformations can be represented by a 3-parameters translation vector T and a 3x3 rotation matrix R that depends likewise on 3 parameters. This transformation can be described using a single matrix **M** equation: $\mathbf{b} = \mathbf{Ma}$, where **a** and **b** are the old and new homogeneous coordinate vectors:

$$\begin{pmatrix} b_x \\ b_y \\ b_z \\ 1 \end{pmatrix} = \begin{pmatrix} & \mathbf{R} & & \mathbf{T} \\ & & & \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_x \\ a_y \\ a_z \\ 1 \end{pmatrix}$$

Several possibilities exist to represent rotation matrices. The most popular representation uses Euler angles corresponding to 3 successive rotations with angles ϕ, θ, ψ around the x,y and z axis, resulting in the following rotation matrix:

$$\mathbf{R} = \begin{pmatrix} \cos\psi\cos\theta & -\sin\psi\cos\theta & \sin\theta \\ \cos\psi\sin\theta\sin\phi + \sin\psi\cos\phi & -\sin\psi\sin\theta\sin\phi + \cos\psi\cos\phi & -\cos\theta\sin\phi \\ -\cos\psi\sin\theta\cos\phi + \sin\psi\sin\phi & \sin\psi\sin\theta\cos\phi + \cos\psi\sin\phi & \cos\theta\cos\phi \end{pmatrix}.$$

The drawbacks of the Euler representation are first, that the matrix coefficients are non-linear and second, that this representation is not differentiable at some singular values (where $\theta = -\pi/2$) [Aya91].

Another well-known representation makes use of unit quaternions that are an extension of complex numbers [FH86]. The unit quaternion is a 4 components vector $\mathbf{q} = [q_0\, q_1\, q_2\, q_3]^t$, where $q_0 \geq 0$, and $q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1$. Each unit quaternion corresponds uniquely to a 3 x 3 rotation matrix :

$$\mathbf{R} = \begin{pmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 - q_0q_3) & 2(q_1q_3 + q_0q_2) \\ 2(q_1q_2 + q_0q_3) & q_0^2 + q_2^2 - q_1^2 - q_3^2 & 2(q_2q_3 - q_0q_1) \\ 2(q_1q_3 - q_0q_2) & 2(q_2q_3 + q_0q_1) & q_0^2 + q_3^2 - q_1^2 - q_2^2 \end{pmatrix}.$$

Unit quaternions have some advantages amongst which simple derivatives and efficient formulations to find rotation matrices directly for point matching problems [FH86, Hor87, BM92].

Another less known representation that is minimal and that equally avoids singularities is the rotation vector $\mathbf{r} = (r_x, r_y, r_z)$ where the direction of **r** is the rotation axis and the norm of **r** is the rotation angle (see details in [Aya91]).

Due two their above mentioned advantages quaternions are used throughout this thesis.

Nonrigid transformations: For the tasks addressed in this work, the transformation **T** has to take deformations into account. The deformations can be global or local, elastic or plastic. The registration may be affine, projective, curved or even completely arbitrary. Figure 4.4 shows how each of the transformations affects an image.

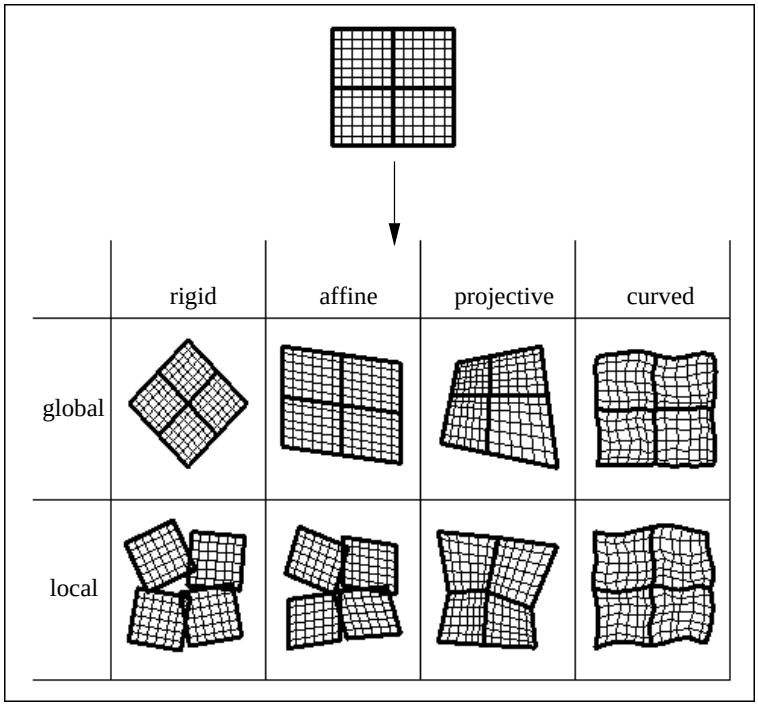


Figure 4.4: Different kinds of deformations

4.2.2 Definition of a disparity function between reference features

Once a relation between Ref_A and Ref_B has been defined, the second step is to extract corresponding features in Ref_A and Ref_B, and to define a disparity function (or a similarity function) between these features. Reference systems Ref_A and Ref_B will be assumed to be registered when the defined disparity function (or similarity function) will be minimal (or maximal). Such an optimization will constitute the last step of the registration process. Obviously, the choice of reference features and of the corresponding optimization method is the core of any registration strategy, most methods differ at this level.

Reference Features

Following the classification proposed in [MV98], reference features can be extrinsic or intrinsic.

One classical example for extrinsic references features are fiducial markers that are attached to preferably rigid structures (bones) of the patient. Subsequent imaging of the patient allows the registration between physical space and the images based on segmentation of the fiducials in the images. This can be very precise but implies an invasive intervention prior to the imaging and real surgery of the patient in order to implant these reference structures. The system for Computer Assisted ACL surgery presented in section 2.2, applies this method.

Intrinsic features can be based on voxel properties, anatomical landmarks or on segmented features like points, curves, surfaces or deformable models (see [Lav96, MV98] for further details).

An interesting concept is to determine virtual points using a functional criterion. This idea has been introduced in [LWM+94] for robot-assisted knee surgery: The center of the femoral head defined on CT images can be registered with the femur's center of motion with respect to the pelvic bone when a passive rotation is applied by the surgeon (or by a robot) during an operation. It also offers interesting options for initializing the rigid parameters of registration algorithms.

Definition of a disparity function

Assuming a set of reference features $\mathcal{F}_\mathrm{A} = \{F_{Ai}, i = 1 \ldots N\}$ has been extracted in Ref_A, with a corresponding set of features $\mathcal{F}_\mathrm{B} = \{F_{Bj}, j = 1 \ldots M\}$ in Ref_B, the disparity function must involve some function of the distances between features F_{Ai} and the features F_{Bj} transformed by the transformation $\mathbf{T}$ to be found.

If features are points, the Euclidean distance is perfect, but for more complex, higher-level features, more sophisticated distance functions have to be used such as the Hausdorff distance for instance which is defined by the maximum of the minimum euclidean distances between each point of a feature and the other features (see also [LCT97]). Most existing methods minimize a weighted least-square criterion of the type

$$D = \sum_i w_i [distance(F_{Ai}, \mathbf{T}(F_{Bi}, t))]^2$$

where w_i are scalar weights for each feature. Nonrigid transformations $\mathbf{T}$ tend to deform a model A until it fits perfectly a model B, thus reducing the disparity function to zero, which is conflicting with the existence of noise and the fact that the nonrigid matching is in general an ill posed problem in general, i.e. the system of equations defining the problem is underdetermined. One solution is to set constraints directly on the transformation $\mathbf{T}$ by limiting the number of degrees of freedom or parameters which are necessary to define an elastic transformation $\mathbf{T}$. For instance, $\mathbf{T}$ can be defined as an adaptive hierarchical spline function [SL96]. Another solution is to set constraints during the optimization process, using regularization theory [Ter88, TF88, Boo89, SL96]. In the latter case, the criterion to be minimized is a weighted sum of the errors between features and a regularization (smoothing) term. Both methods are used in this work.

4.2.3 Robustness

As said before, most methods minimize a least square criterion. This is appropriate if the noise in the data has a Gaussian distribution (Maximum-likelihood estimation). Noise is introduced by sensors, calibration errors, data approximation through inexact representations, errors of the feature extraction procedures and phenomena not taken into account in the model of the transformation.

In many applications, visual data are not only noisy, but also contain *outliers*, data that are in gross disagreement with a postulated model. Outliers, which are inevitably included in an initial fit, can so distort a fitting process that the fitted parameters become arbitrary. A similar situation occures when there is only a partial overlap between the model and the data.

In such circumstances, *robust* estimation methods [HRRS86, Hub81] can be applied. Robust methods aim at continuing to recover meaningful parameters even when the data contain spurious data. Elementary procedures imply iteratively repeated rejection of data above a given threshold [LS95] (Regression Diagnostics) or using only a pre-defined fraction of the best data (for instance 80%) [BLS93, JRH92, HR93].

The least median of squares (LMS) method [Rou84] minimizes the square (or absolute value) of the median residual for the data. This ignores the largest residuals in the sample, therefore this method is robust in the presence of outliers. Since sorting to find the median becomes expensive if the data set is very large, usually random subsets are considered in this case.

So called M-estimators (that maximize the likelihood for a particular error distribution) are an other class of robust estimators. A robust estimator based on the least square error norm $\rho(x) = x^2$ has the advantage of being convex near the minimum, but gives unbounded influence as the data error goes to infinity. More precisely this influence is proportional to the derivative ψ (called influence function) of ρ. For the least square case $\psi(x) = 2x$. To increase the robustness, one can think of error norms that are more forgiving about outlying measurements, i.e. increasing less rapidly than x^2. Some of the more commonly M-estimators use the following error norms ρ:

$$\text{Tukey:} \qquad \rho(x,k) = \begin{cases} \frac{c^2}{6}(1-(1-(\frac{x}{c})^2)^3) & |x| \le c \\ \frac{c^2}{6} & |x| > c \end{cases}$$

$$\text{Geman-McClure:} \quad \frac{x^2/2}{1+x^2}$$

In [ea99c] for instance authors use robust estimators for single modality image registration based on optical flow. In [Nik98] authors use the Geman-McClure error norm for registration of multi modality images. In [MEF99] the Tukey error norm is used to reject outliers in surface based registration methods for orthopedic surgery.

4.2.4 Optimization of the disparity function

The defined disparity function has to be minimized using a dedicated optimization procedure. Some possible criteria for classification of optimization methods are given in the following.

Global versus local matching: First, one can distinguish two main categories: In *global* methods, only one criterion depends on the searched parameters. This criterion takes all features of the model space and all features of the data space into account. In *local* methods, features of the model space are matched individually with features of the data space, and each matched pair of features gives directly some possible parameters of the transformation $\mathbf{T}$, or at least some constraints on these parameters.

Nonlinear optimization: Very few methods give a direct solution: this is only the case for rigid registration of points or planes. Therefore, minimization techniques often require nonlinear optimization using iterative procedures. If the disparity function does not allow to compute the derivatives in a reasonable time (analytically or numerically), algorithms using only function evaluations have to be applied. Otherwise more powerful techniques using first or second derivatives can be used. As it will be seen in the subsequent chapters, for the registration methods proposed in this dissertation, the partial derivatives with respect to the disparity function can be calculated analytically, thus gradient descent techniques can be applied.

Local minima: When iterative optimization is used, there is always a risk that the method fails into a local minimum, raising the issue of *initialization* and *interaction*. Focusing first on rigid-body registration, two kinds of local minima can be distinguished qualitatively. A first class contains a series of many local minima spread in a small neighborhood of the true global minimum. These minima occur when registration is performed between two surfaces that differ only in fine details. They can be avoided by an appropriate discretization or data smoothing. A second category includes large local minima which are quite far away from the true global minimum. Their existence, location and number is very shape-dependent. However, for application in the medical field, there is very often some a priori knowledge about the region where the minimum has to be searched. This is due to the fact that physicians, surgeons and radiologists already use a reference system, which is the patient reference system (see appendix B.1). Typically, the rotation matrix between the data space and the model space is known with uncertainties of about $\pm 30°$ around each axis. For translation components, using an initial translation superimposing the feature centroids of both spaces is usually suitable. In this work simple interactive pre-alignment is used. In [Rot00] some more sophisticated initialization methods are discussed in more detail. For *nonrigid* registration where the parameter space has a much higher dimension, local minima are more likely to occur. In this case, it is possible to use a stochastic method (e.g. based on genetic algorithms as used in [JR93] for instance) in order to find the global minimum. Multi-resolution algorithms also provide an efficient solution: A well known implementation is presented in [BK89a] for elastic volume registration. In this work the later approach is applied in chapter 5.

Figure 4.5 summarizes some of the given classification criteria for registration methods of medical images [MV98].

4.3 Deformable models

In recent years there has been considerable interest in methods that use deformable models, or atlases, to interpret images. Although originally developed for application to prob-

Dimensionality				
Spatial dimensions only	Time series (spatial dimensions + time)			
2D/2D	2D/2D			
2D/3D	2D/3D			
3D/3D	3D/3D			
Nature of registration basis				
Extrinsic		Intrinsic	Non-image based (calibrated coordinate systems)	
Invasive		Landmark based		
Stereotactic frame		Anatomical		
fiducials (screw markers)		Geometrical		
Non-invasive		Segmentaion based		
Mould, frame, dental adapter, etc.		Rigid models (points, curves, surfaces)		
fiducials (skin markers)				
		Deformable models		
		Voxel property based		
		reduction to sclars/vectors (moments, principal axes)		
		Using full image content		
Nature of transformation				
rigid	Affine	Projective	Curved	
Domain of transformation				
Local	Global			
Interaction				
Interactive		Semi-automatic	Automatic	
Initialization supplied		User intializing		
No initialization supplied		User steering/corrrecting		
		Both		
Optimization procedure				
Parameters computed		Parameters searched for		
Modalities involved				
Mono-modal	Multi-modal	Modality to model	Patient to modality	
CT	CT-MR	CT	CT	
MR	US-CT	MR	MR	
US	X-ray CT	X-ray	X-ray	
X-ray	etc.	etc.	etc.	
etc.				
Subject				
Intrasubject	Intersubject	Atlas		
Object				
Head	Limbs (orthopedic)	Spine and vertebrae	etc.	

Figure 4.5: Classification of registration methods

lems in computer vision and computer graphics, the potential of deformable models for use in medical image analysis has been quickly realized. The applications include segmentation, registration, shape analysis, motion tracking etc. The inherent (forced) continuity and smoothness of a model can compensate for noise, gaps and other irregularities in object boundaries. Furthermore, the representation of a model can provide a compact description of object shape. These properties lead to a robust and elegant technique for lining sparse or noisy local image features into a coherent and consistent model of the object. A deformable-model based segmentation scheme, used in connection with image preprocessing, can overcome many of the limitations of manual slice editing and traditional image processing techniques. Connected and continuous geometric models consider an object boundary as a whole and can make use of a priori knowledge to constrain the segmentation problem. Deformable models can be used for matching anatomic structures by exploiting (bottom-up) constraints derived from the image data together with (top-down) *a priori* knowledge about the location, shape, etc. of these structures. For a general overview about deformable models see for instance [MT96, Lel99]. In [MDSA00] authors discuss the different possible representations for deformable surfaces.

The different anatomical modeling methods described in the literature can be placed mainly at the top three levels of the image interpretation pyramid in figure 4.1 depending on the amount of a priori knowledge incorporated into the model.

4.3.1 Incorporating *a priori* knowledge

The need to incorporate prior knowledge into image segmentation methods is nowadays widely recognized especially in medical imaging, where many aspects of the imaging conditions are difficult to control. The incorporation of knowledge about the shape, location, orientation, appearance and spatial context of an organ is essential and may be incorporated into a deformable model in the form of initial conditions, data constraints, constraints on the model shape parameters or into the model fitting procedure. The use of implicit or explicit anatomical knowledge to guide shape recovery is especially important for robust and automatic interpretation of medical images. It is essential to have a model that not only describes the

- shape (smoothness, variability, topology, etc.)
- size
- location, orientation
- boundary
- texture (for volumetric models only)

of the target object but that also permits expected variations in these characteristics. The knowledge can be gathered by different means; regarding the shape it can be for instance physically or population based.

Automatic interpretation of medical images can relieve clinicians of the time consuming aspects of their work, while increasing the accuracy, consistency and reproducibility of the interpretations. As the task addressed in this work does not only consist in *interpolating* but especially in *extrapolating* data, incorporating *a priori* knowledge about the shape to be recovered in a suitable manner, is one of the key problems to solve.

As pointed out in [MT96] it is difficult to classify deformable models, due to the variety of different approaches. According to the requirements in this work, the following criteria are important:

- **Automation, accuracy and robustness:** High automation is always the long term goal for image analysis systems but often only semi-interactive tools provide robust and accurate results in existing systems. Actually, interactivity can be considered as a basic way to introduce a priori knowledge in the segmentation process, i.e. the anatomical knowledge of the expert.

- **Generality versus specificity:** Generality is the basis of deformable model formulations with local shape parameters such as snakes (see 4.3.2). Alternatively, highly specific or constrained deformable models exist. The objective of this work is to investigate a general method to build specific models for each object-class (femur, vertebra, etc.) to be reconstructed.

- **Surface versus solid models:** Depending on the task, either surface models or volumetric models can be more appropriate.

- **Geometric Nature:** One can distinguish between the shape description and the deformation description. The shape description may confine the model to represent shapes of restricted or not restricted topology (e.g. fourier descriptors), and may or may not allow a topology change throughout the evolution of the model. Within this dissertation the goal is to recover the shape of an object with known, stable but possibly complex topology as a vertebra for instance. The complexity of possible model deformation may be unrelated to the complexity of the model shape. A high resolution surface mesh for instance may be restricted to undergo simple affine transformation only.

4.3.2 Popular Deformable Models

In the following section some of the more popular deformable models are briefly presented.

Snakes

A widely acknowledged object representation applied for segmentation of medical images was described by Kass [KWT88], who introduced the active contours, also referred to as snakes. This flexible contour is an energy minimizing spline curve with associated stiffness and elasticity. A parametric snake is a curve expressed in coordinate functions($x(s)$ and $y(s)$); where s represents the parametric domain (0,1). The shape of the contour is governed by an energy functional:

$$E(v) = \int_0^1 \left(w_1(s) \left| \frac{\partial v}{\partial s} \right| + w_2(s) \left| \frac{\partial^2 v}{\partial s^2} \right| \right) ds + \int_0^1 P(v(s)) ds + \int_0^1 E_{constr.} ds$$

where the first integral term represents an internal deformation energy of the model, which is balanced with an external scalar field $P(v)$, typically defined from an image feature, such as the local image gradient. Parameter functions $w_1(s)$ and $w_2(s)$ represent two physical properties of the contour, i.e. the ability to stretch and bend respectively. These functions can be used to impose a preferred shape on the model and to locally control shape characteristics like the object smoothness of the resulting segmentation. The third term E represents shape constraints introduced by the user, e.g. by fixating a point of the contour to an image point. The final shape of an active contour in an image corresponds to a minimum in $E(v)$, which can be found by numerically solving the Euler-Lagrange equation. Extensions of the 2D snake model to 3 dimensions (deformable balloons) have since been reported, as well as many modifications of the original energy formulation to improve robustness of the snake and balloon methods with respect to spurious feature points and initial positioning, transitions in topology and simultaneous detection of multiple objects.

The drawback of such models is that they are free to take almost any smooth shape, and are therefore *non-specific*. That is, they can produce examples of the object of interest which are outside the normal variation of shape for that object.

"Hand Crafted" models

Flexible models can be built up from simple subcomponents, such as circles, lines, or arcs, which are allowed some degree of freedom to move around relative to one another, and possibly change scale and orientation. In [YCH92] for instance parts of the face, such as the eyes and mouth, are modeled in this way. Lipson *et al.* [ea90] apply a similar scheme to map elliptical models of vertebrae onto CT images of the spine.

Superquadrics

Deformable models based on superquadrics are one example of deformable models incorporating *a priori* knowledge by applying constraints on global shape parameters. Superquadrics form a family of implicit surfaces obtained by extension of usual quadrics. They are obtained by the spherical product [Bar81] of two 2-D curves. Their advantage is their capability of covering a wide variety of shapes with only a small number of parameters.

However, if superquadric shapes give a good approximation of a shape, they are not sufficient to describe more complex (anatomic) surfaces. To overcome this restriction, they can be coupled with local shape parameters. Bardinet *et al.* [BCA94] for instance fit a deformable superquadric to 3D cardiac data and refine the superquadric fit using a volumetric deformation technique known as free-form deformations (FFDs) [SP86].

Finite Element Models

Finite element methods can be used to model variable image objects as physical entities with internal stiffness and elasticity. Pentland and Sclaroff [PS91] use three-dimensional models which act like lumps of elastic clay. They derive modes of vibration of a suitable base shape, such as an ellipsoid, and build up shapes from different modes of vibration. The first modes are large-scale variations of shape; the higher order modes are more localized. They fit models to range data by an iterative process. Terzopoulos and Metaxas [TM91] present a similar idea using deformable superquadrics (see above). Nastar and Ayache [NA92] apply a finite element approach using the vibrational modes of an example of the shape to be modeled.

Volumetric models

Volumetric templates consist of a segmented voxel set and are matched to image data by deforming the model on the basis of attraction forces, which are derived from local similarity measures. The dimensionality of the defined transformations determines the matching accuracy that can be achieved with these deformations. Several approaches have been described based on elastic [BK89b, CRM94] and thin-plate spline interpolants spanned by landmarks [ECNM91] for instance. Such models are most commonly applied to segmentation of brain structures. Due to the locally distributed nature of shape knowledge in these models, the model-image matching is computationally an order of magnitude more expensive than for boundary template approaches described so far. Since volumetric templates are matched as a whole, the topological structure of the model is preserved throughout the matching procedure, which makes it suitable to segment multiple objects in a scene simultaneously. Therefore, these models can be placed on the scene interpretation level in the image interpretation pyramid in figure 4.1.

All mentioned methods have the advantage that the models are relatively easy to construct and allow a compact parametric representation of a family of shapes, but they do not represent the actual variation of shape within a (natural) population. Therefore, investigations have been made to develop models which can only deform in ways which are characteristic of the objects they represent. To limit the number of shape parameters, it is necessary to find (uncorrelated) model parameters describing typical modes of shape variation.

4.3.3 Statistical shape models

Fourier Series Shape Models

Different population based models are known from the literature. One possibility is to use statistical models based on Fourier representations such as [SD92, SKBG96].

In [SD92], a method of modeling shapes by an expansion of trigonomic functions is proposed. The authors derive distributions for each of the parameters over a training set, and while fitting the model to an image, maximize a probability measure determining

how likely it is that the current example is the desired object. Trigonomic basis functions are not suitable for describing general shape; for example, using a finite number of terms, they can only approximate a square corner. The relationship between variations in shape and variations in the parameters of the trigonometric expansion is not straightforward.

Volumetric models

A second important statistical shape modeling method for segmentation purposes is based on spatial normalization of a set of voxel maps. By optimally registering a set of segmented volumes into a standardized space by applying affine or higher dimensional transformations, an image scene can be expressed as a probability map or as an average shape with a locally defined variance measure respectively. These models are generally applied to segmentation problems by weighting a feature-based probability density function with a spatial probability distribution of an organ shape in a Bayesian formulation (see also section 4.3.2. These types of models are commonly applied for brain warping.

Another approach is to consider a model with modal representation based on statistical analysis directly applied to the nodal representation of a mean contour [CTCG95a]. This thesis is based on the latter approach which is therefore explained in further details in the next section.

The Statistical Shape model of Cootes and Taylor

A widely acknowledged statistical shape model is the Point Distribution Model (PDM) as introduced by Cootes et al. [TCCG92] A PDM describes the average shape and characteristic shape variations of a set of training samples, which are given in the form of a set of points on the sample boundaries. principal component analysis (PCA), also known as Karhunen-Loeve transformation is used to decompose the model in eigenmodes obtained from the shape training set. The decomposition basis size depends on the number of shapes contained in the training set. Since only shapes composed by a linear combination of the eigenmodes can be represented, PCA restricts the shape variation of the model close to the known statistical shape variation.

Given a collection of N 3D training shapes of an object, the Cartesian coordinates of M corresponding points are recorded for each image. Each training example is represented by a vector $\mathbf{m} = (x_1, y_1, z_1, ..., x_M, y_M, z_M)$. After alignment of the training shapes the pointwise mean shape

$$\bar{\mathbf{m}} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{m}_i \tag{4.1}$$

is then calculated. Modes of variation are found by applying principal component analysis to the data. That is the modes are represented by the $3M$ orthonormal eigenvectors $\mathbf{e}_i$ of the covariance matrix describing the deviations of the training shapes from the mean shape:

$$\mathbf{C} = \frac{1}{N-1}\sum_{i=1}^{N}(\mathbf{m}_i - \bar{\mathbf{m}})(\mathbf{m}_i - \bar{\mathbf{m}})^T$$

A new instance of the shape is generated by adding linear combinations of the t most significant variation vectors to the mean shape

$$\mathbf{m} = \bar{\mathbf{m}} + \sum_{i=1}^{t} w_i \mathbf{e}_i \tag{4.2}$$

where w_i is a weighting for the i^{th} variation vector. PCA is based on the hypothesis that the data correspond to a single multidimensional Gaussian distribution, thus justifying a linear representation. Vector $\mathbf{m}_i$ varies inside an N dimensional hyper ellipsoid. The eigenvectors of $\mathbf{C}$ represent the hyper ellipsoid axis directions while the eigenvalues are the axis amplitudes. By applying limits to the variation of w_i, for instance $|w_i| \leq \pm 3\sqrt{\lambda_i}$, it can be prevented that a new shape instance varies too much from the mean shape.

Usually, N is much lower than $3M$ and diagonalization of the covariance matrix results in N eigenvectors associated to N non zero eigenvalues. It can be shown that the eigenvectors of the covariance matrix can be calculated from a smaller $N \times N$ matrix derived from the same data in this case [CTCG95b, TP91]. Because the eigenvector calculation time is proportional to the cube of the size of the matrix, this can give substantial time savings, as $3M$ may be in the range of several thousands, while N normally is much smaller. The number of usable modes t depends on the size of the training set. It must be large enough compared to the parameter vector size $3M$ to allow the model to recover an object.

One important difference between matching methods based on such a PDM and the matching mechanism for snake models for instance is the absence of an energy functional based on elastic material properties. For PDMs, the image matching is performed by calculating a suggested boundary location for each point in the PDM, based on image information. This allows an elegant coupling between high-level knowledge about object shapes and low-level image features. The model pose and shape-parameters are iteratively updated to optimally fit the hypothesized shape, where the model is only allowed to deform along the most characteristic eigen deformations.

Because statistical shape models allow a coupling between low-level image data and higher level knowledge about individual organ shape and their spatial context in a scene, they correspond to the scene interpretation level in the image interpretation hierarchy in figure 4.1.

4.3.4 Shape Analysis

The construction of a training set implies to define M corresponding points on a set of N shape instances. In two dimensions this point correspondence is often defined manually. This is fastidious and may become easily impractible in the 3D case. Indeed, determining corresponding points on smooth surfaces is a difficult task for a human operator except for feature points such as curvature extrema. Such visually determinable homologous points are called anatomical landmarks.

These landmarks are usually sparse (depending on the complexity of the shape) and thus not sufficient to build a PDM upon; A PDM requires a *dense* set of corresponding points, which are not necessarily determinable on a certain specimen alone but that correspond across all the specimens of a data set under a reasonable model of deformation from their common mean. Thus they can be used as if they were anatomical landmarks and are therefore called semi or pseudo landmarks [Boo96]. Some authors propose methods to define dense corresponding point sets based on an application dependent assumption, often necessitating very specific acquisition protocols. The development of more generic methods to define point correspondence for two and especially three dimensions is currently an active field of research. Most known methods from the literature rely on nonrigid registration of deformable models and assume that the training objects have been already segmented. One class of methods relies on a known sparse set of landmarks; a dense set of semi-landmarks is then defined by registering the corresponding landmark set applying a specific deformation. In [STA96] a set of anatomical landmarks called crest-lines are computed to perform shape analysis on a skull population. Crest-lines are also used in [ea00b] to establish initial correspondence. A dense set of corresponding points is then computed by applying a diffusion algorithm. In [LK00] initial correspondence is based on a manually defined set of sparse landmarks on each object. A dense mapping is then computed using thin-plate interpolation. To achieve a final precise mapping between the two object boundaries and to regularize the surface mesh, a mesh relaxation is applied subsequently. In [BT99] two shapes are represented by dense triangle meshes which are matched via sparse triangulations obtained by triangle decimation from the dense mesh. The matching is based on a symmetrical version of the iterative closest point algorithm (see section 4.4). To overcome the problems of possible surface folding related to this approach, the same authors propose in [BT00] to transform the surface to a planar domain by means of harmonic maps. This approach is only applicable to single part shapes that are topologically isomorphic to a disk. In [ea99a] Kelemen et al. use Fourier-descriptors to automate the landmark generation. As mentioned in section 4.3.3 models based on Fourier-descriptors likewise suffer from topological restrictions. Kotcheff and Taylor [KT98] use direct optimization to place landmarks on sets of closed 2D curves. Though good results are reported, possible extension to three dimensions is questionable due to the extremely high computational costs for the applied optimization method based on a genetic algorithm. Caunce and Taylor [CT98] describe a method to match sulcal fissures using a modified ICP algorithm; the landmarks are progressively improved by adding more structural information into the matching process. However, an extension of this application specific method for a general use outside the brain does not seem straightforward. Wang et al. [ea00c] use a surface registration technique to find 3D point correspondences taking curvature information into account. Authors suggest to use their method to build 3D point distribution models but no results are reported. In [ea01c] authors propose a method for building statistical shape models, where no segmentation of the training shapes prior to the shape matching is required. A free-form elastic registration technique based on maximization of normalized mutual information is used to establish correspondence between images. As proposed in [FL98] subsequent principal component analysis is applied directly to the deformation lattice instead of to the semi-landmarks. For more fundamental work on shape analysis

and morphometrics the reader is referred to [Boo96, GM98, Lon98].

In the following chapter, a method which performs an automatic landmark point generation while ensuring point correspondence between the training shapes is described.

4.4 Iterative Corresponding Point Matching

In many cases the registration problem can be formulated as an *iterative corresponding point matching problem*. Starting from an initial position and orientation, during the first step of each iteration corresponding points in the data space and model space are determined. During the second step a new transformation is computed by applying one iteration of the chosen optimization method. The result is applied to the data space and the whole procedure is iterated until convergence.

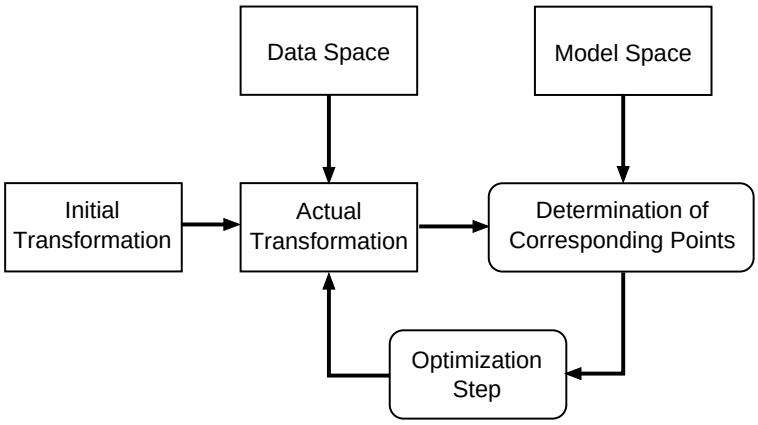


Figure 4.6: Flow diagram for Iterative Corresponding Point Matching

The rigid and non rigid parameters (if existing) of the transformation can be coupled and optimized together or treated separately. The optimal rigid registration between two corresponding points is a well-known problem for which a direct method exists in the 3D case. It is based on quaternions and its mathematical description can be found in appendix C. If the correspondence between the data set and the model set is defined by simply assigning the closest model point to each data point (using Euclidean distance measure) the above described iterative procedure is well known as the Iterative Closest Point algorithm [BM92]. It is possible to speed up this algorithm considerably by extrapolating the last 3 quaternions [BM92]. See for instance [SHK95] for further improvements. The ICP can also be applied to surface registration. For intra-operative application it is mandatory to find nearest points on the model surface very quickly, which is not obvious for complex shapes. One possibility to speed up the search is to use k-d-trees, a form of hierarchical space-decomposition, which is discussed further in section 6.4. Of course more complicated correspondences can be defined, thus allowing registration of a 3D model with 3D gray level images (chapter 5), registration of a 3D model with sparse 3D point data (chapter 6) and registration of a 3D model with 2D images (chapter 7).

4.5 Conclusion

Deformable models are very interesting for computer vision problems such as segmentation and registration due to their capability of exploiting (bottom-up) constraints derived from the image data together with (top-down) a priori knowledge about the structure to be analyzed. For problems such as the ones addressed in this thesis, where only incomplete information about the object is available (sparse scattered point data, few X-ray images) to uniquely determine the solution, models that are limited to deform in ways which are characteristic of the object they represent seem most promising. Models that do not limit their shape to the actual variation of the target object within a natural population, might be capable of fitting the available data very well, but are not sufficiently constrained in regions without data.

The statistical shape model introduced by Cootes and Taylor describes the average shape and characteristic shape variation of a set of training samples that are defined by a corresponding set of boundary points. Along with its simple representation, the capability of dealing with aribitrary shape topologies favours the use of this model as the objective is to develop a generally applicable concept for any specific anatomical shape.

Figure 4.7 shows a flow diagram of the chosen global approach whose different components are presented in the next three chapters. Chapter 5 presents a method for automatic extraction of training shapes from unsegmented CT images using a generic model. Chapter 6 presents a method to obtain training shapes if shape data is available in the form of unorganized sparse surface point sets (gathered by manual digitization for instance). The subsequently computed statistical shape model can be matched to sparse scattered point data in order to obtain a patient specific model (nonrigid 3D/3D registration), as shown in the same chapter. Chapter 7 explains how to obtain such a patient specific model by fitting the statistical model to a small set of calibrated X-ray images (nonrigid 3D/2D registration). It is also demonstrated that the combination of both methods resulting in a hybrid registration is straightforward.

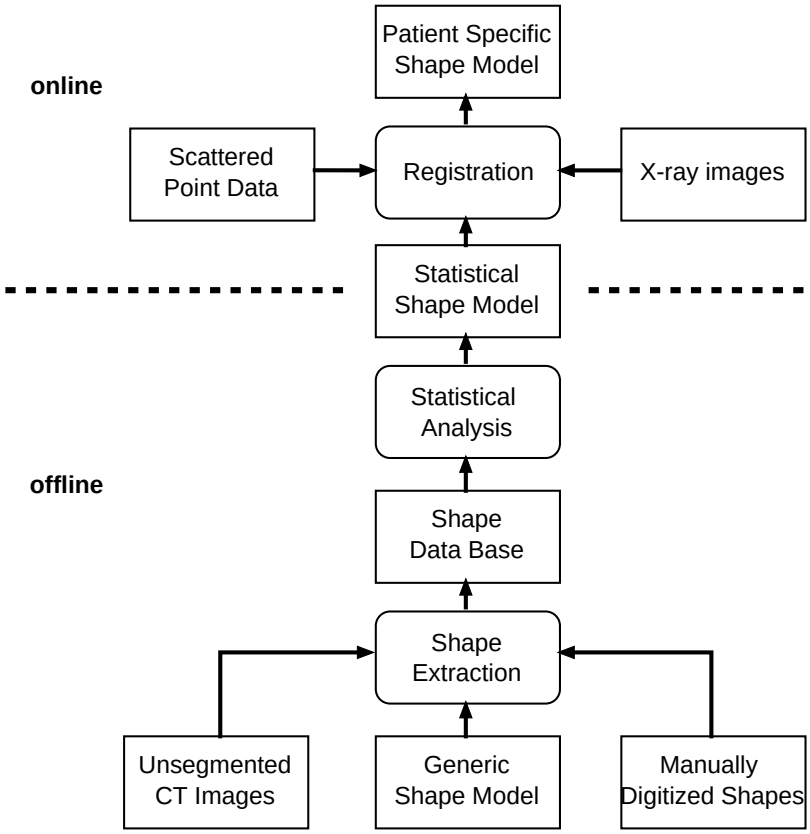


Figure 4.7: Flow diagram for Shape Recovery by Nonrigid Registration of a Statistical Shape Model with intra-operative data. The statistical shape model is computed offline and can be matched intra-operatively with scattered point data or X-ray images. It is also possible to use both type of data simultaneously (hybrid matching)

Chapter 5

Non-Rigid Registration of a generic deformable surface model to gray level images: Application to automatic shape extraction from CT images

As discussed in section 4.3.4 manual determination of a dense set of corresponding 3D points in a group of training shapes are practically infeasible. Most known methods from the literature aiming to automate this process rely on binary images.The necessary prior segmentation is often done (semi) manually, i.e the segmentation and registration of the training shapes is usually performed in separate independent steps. This chapter presents an approach trying to perform both steps simultaneously, thus allowing to take into account existing duality between segmentation and registration (see section 4) that is neglected in most methods described in section 4.3.4. As discussed in section 4.1.2, unsupervised segmentation based on purely local operators can be very difficult,while adequate *a priori* knowledge contained in a deformable model can make the segmentation process easier and more robust. In the case of vertebra segmentation in CT-images for instance, using standard low level segmentation tools poses the following problems:

- It is difficult to isolate one vertebra from the others, because of the limited spatial resolution of the 3D scalar field and the extremely narrow bone gaps at the articulating surfaces (see Figure 5.1)
- It is difficult to differentiate the external cortical surface from the internal cortical surface.

The primary objective of the method presented here is the automatic shape extraction of bones in CT images. Based on a general shape template a representation of the external cortical surface is inferred from the model to the data providing a dense set of corresponding points between the template and the data, which can be used for subsequent statistical analysis. It is worth to note that the matching will be based only on

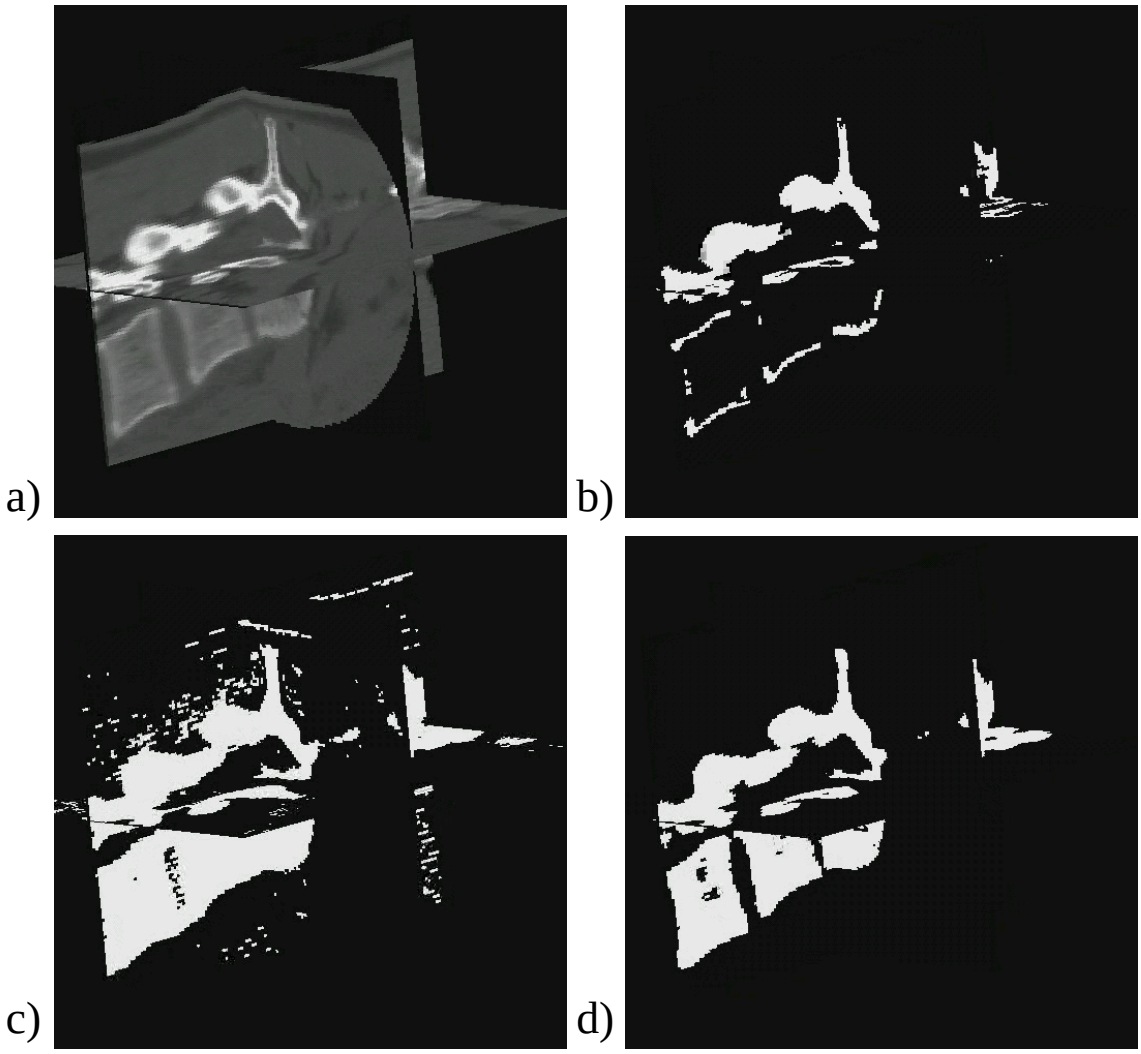


Figure 5.1: Thresholding does not allow to separate a single vertebrae in a CT image: Gray level image (a), high threshold (b), low threshold (c), intermediate threshold (d)

object boundaries, in contrast to the work in [ea01a] for instance, where the gray level information of the whole volume is used to drive the deformation. However the resulting deformation will be volumetric thus allowing for automatic inference of other structures embedded in the volume.

5.1 Approach

The method performs a least square minimization of the distances between model boundary points and matched feature points. The deformation defined by the computed displacement field are described as a warping of the space containing the surface model, based on 3D tensor-product deformation splines. For increased efficiency, these so-called *Free-Form Deformations* [SP86] are applied in a multi-resolution framework. The result is a rapid and efficient registration algorithm which does not require the prior segmentation (manual or automatic) of features in the data image, and which can work on arbitrarily shaped surfaces. The proposed method is conceptually following [SL96, LBC+99] but introducing important improvements and extensions sharing ideas from [ea99b, Pic97, CMT98, ea00a]. The problem can be formulated as a minimization of a cost function

$$E(\mathbf{p}) = \sum_{i=1}^{N} [dist(P_i, T_{\mathbf{p}}(M_i))]^2 + \mathcal{R}(\mathbf{p}), \tag{5.1}$$

where $dist$ is the euclidean distance between a model point M_i and its corresponding data point P_i in the gray level image. T is a suitable deformation function depending on the parameter vector $\mathbf{p}$ and where $\mathcal{R}$ defines a regularization term which is applied to T in order to smooth the deformation.

To solve the minimization problem, four components are required:

- Generation of a suitable generic shape template
- A strategy to define a correspondence between the model points and the image data
- A suitable representation for the geometric transformation describing the deformation D.
- An iterative minimization algorithm

The data flow diagram of the proposed approach is shown in figure 5.2.

5.2 Construction of a generic Shape Model

In the following subsections the different necessary steps for the construction of a generic shape template are detailed for the example case of a lumbar vertebra.

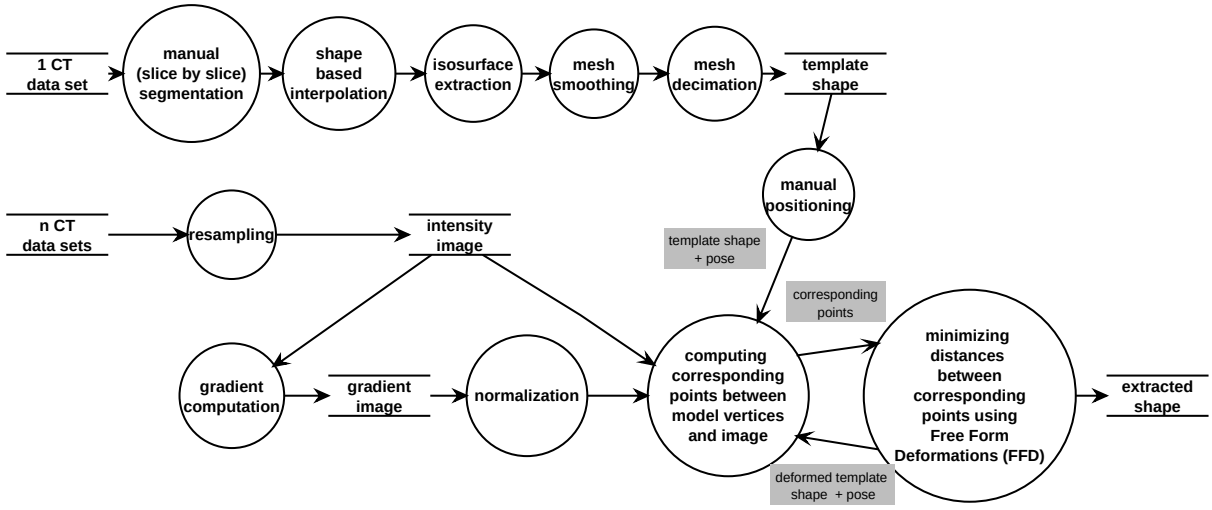


Figure 5.2: Data Flow Diagram of the proposed method for automatic shape extraction from CT images

5.2.1 Manual slice based segmentation

The template model is based on a single CT scan of a non pathologic lumbar spine. The vertebra is segmented using conventional two-dimensional interactive segmentation techniques resulting in a binary volume.

5.2.2 Shape based interpolation

Usually CT voxel data are not isotropic i.e. the out of plane resolution is much lower than the in plane resolution. Therefore slice interpolation becomes necessary. Interpolation techniques may be broadly divided into two categories [GU97]: scene-based and object-based. In scene-based methods, interpolated values are determined directly from the density values of the given scene. Object-based methods such as shape based interpolation, on the other hand, extract some object information from the original gray level data and then use this information to guide the interpolation process. Due to their better performance ([GU97], see Figure 5.3) a shape-based contour interpolation method, based on a distance transformation of the binary volume has been implemented.

5.2.3 Isosurface Extraction

In order to obtain a mesh-based boundary representation of the segmented vertebra from the resampled interpolated volume the marching cube algorithm [Lor87] is applied. The extracted isosurface $f(x, y, z) = c$ contains large sets of triangles even in regions where the curvature is very low, see Figure 5.5 b.

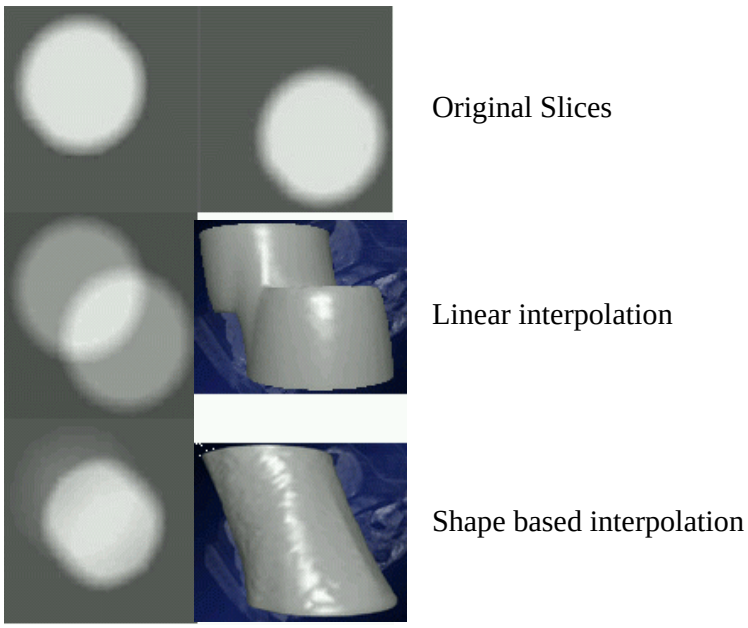


Figure 5.3: Better performance of shape based interpolation (bottom) when compared to linear interpolation (middle).

5.2.4 Model simplification

A smoothing and triangle decimation algorithm is therefore subsequently used to reduce the number of triangular surface elements in these model regions. Models of different resolutions can be generated according to the accuracy requirements. A highly decimated model can be fitted faster to the data. As mentioned before, the deformation is applied to the volume in which the generic model is embedded; therefore the resolution of the model can be altered without interfering with the deformation process.

Decimation and smoothing are two distinct operations that can be performed on the model. With decimation, triangles are selectively removed from the model. With smoothing, vertices of the existing triangles are adjusted in order to reduce the curvature variations of the model. Assuming that vertex points are adjusted along the direction of their surface normals to establish correspondence with the data (see section 5.3), it is important that the model is smoother than the data (Figure 5.4 a). A so-called overfitted model will otherwise lead to wrong correspondences (Figure 5.4 b).

Mesh smoothing

The overall effect of smoothing is reduced high frequency surface information. During smoothing only the geometry of the data set is modified, the topology remains the same. The most common method for performing smoothing is called Laplacian smoothing and consists in replacing the value of a vertex location with the mean of its adjacent vertices. This simple algorithm will tend to flatten the surface and shrink the object, see Figure 5.5 d. To overcome these problems, a variation of the Laplacian smoothing method is

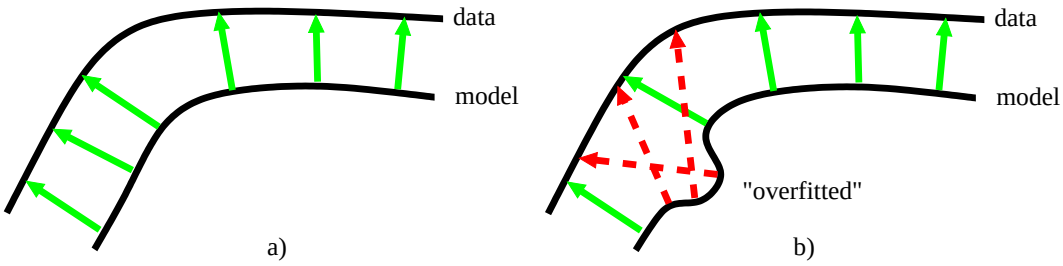


Figure 5.4: Arising problems if the generic model is overfitted

therefore used ([VMH99]) that reduces the effect of shrinking, while preserving the effect of smoothing, see Figure 5.5 e.

Mesh Simplification/Decimation

Polygon reduction algorithms simplify the mesh depending on local curvature, they are therefore most efficient when applied after the mesh smoothing. Various techniques exist, based on vertex decimation, vertex clustering, simplification envelopes or wavelets for instance (see [HG97] for a survey). The method used in this work is based on edge contraction [GH97]. The algorithm takes the two endpoints of the target edge, moves them to the same position, links all the incident edges to one of the vertices, deletes the other vertex, and removes any faces that have degenerated into lines or points. Typically, this removes two triangular faces per edge contraction. The result of this operation is shown in Figure 5.5 f.

5.3 Establishing Correspondences

In [SL96] a 3D Canny-Deriche filter to compute image gradients at each voxel location is applied to the model and the data. Non-maxima suppression processing is then performed to extract the local maxima of gradients on each 3D image. Hysteresis thresholding [HM93] is finally applied on both images of local maxima to retain only a reasonable set of features. The threshold values for the data image are chosen such that the resulting image is oversegmented with respect to the model image, thus increasing the likelihood that a model point can be matched with a data point.

Correspondence between the model and the data is then established by assigning the closest data point to a model point (based on a Euclidean distance measure) using 3D precomputed distance maps. This can easily lead to wrong correspondences as shown in Figure 5.6 a, for the case that extremities of thin parts of the model are not attracted by extremities of the data. Another case may occur when internal surfaces are attracted by external surfaces.

To cope with these issues, the same authors extended the method in [LBC+99] by using a 6D distance function as proposed by Feldmar et al. [FA96]. This distance is defined between a 6D data point $D_j = (x_{D_j}, y_{D_j}, z_{D_j}, Gx_{D_j}, Gy_{D_j}, Gz_{D_j})$ and a 6D model point

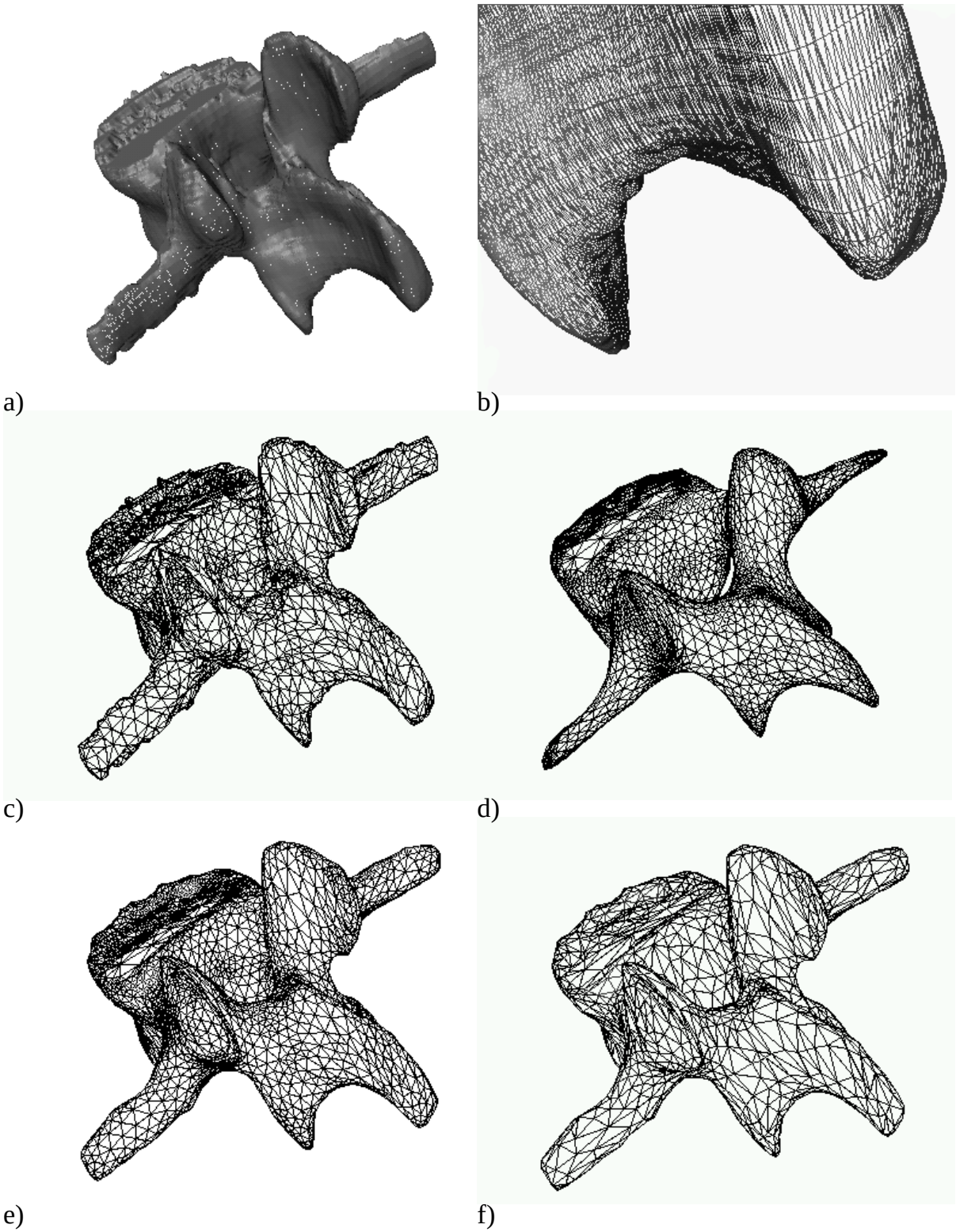


Figure 5.5: a) marching cube high resolution mesh (257914 facets), b) close up of high resolution mesh, c) decimated mesh (10000 facets), d) smoothed mesh using laplacian filtering, e) smoothed mesh using improved laplacian filtering, f) further reduced mesh (4000 facets)

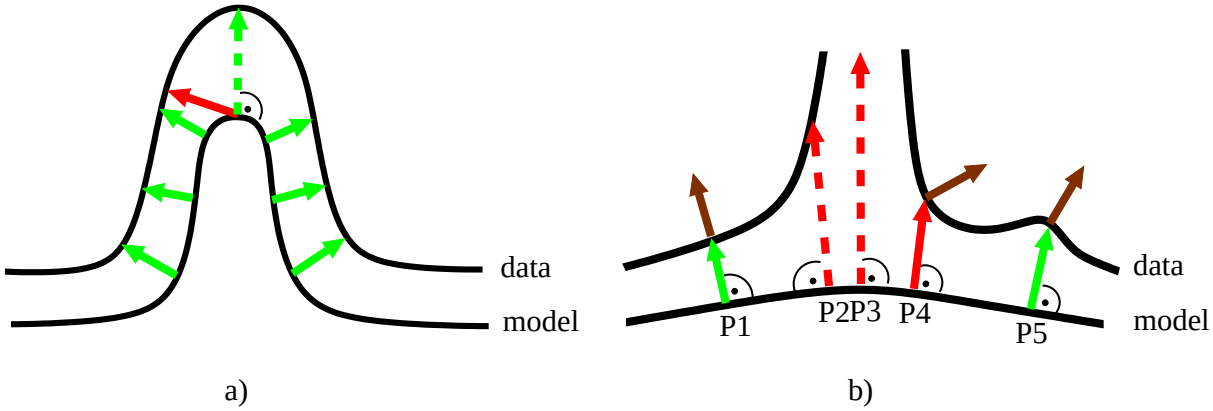


Figure 5.6: Wrong correspondences with the closest point approach (a). The correspondence depicted by the dark arrow is obviously wrong; searching for a corresponding point in the direction of the surface normal (dashed arrow) yields a better result. Five different vertex matching scenarios are shown in (b). Only points P1 and P5 are considered valid matches because (1) the angle between the model surface normal and the direction of the gradient in the image is lower than Δ_{angle}, and (2) the distance between the model point and the matching point in the data is lower than Δ_{range}.

$M_i' = (x_{M_i}, y_{M_i}, z_{M_i}, Gx_{M_i}, Gy_{M_i}, Gz_{M_i})$ constituted of 3D positions of gradient points and 3D coordinates of gradients by the following expression :

$$\begin{aligned} & d^2_{6D}(D_j, M_i') = \\ & (x_{D_j} - (x_{M_j})^2 \quad +(y_{D_j} - (y_{M_j})^2 \quad +(z_{D_j} - (z_{M_j})^2 + \\ & \alpha(Gx_{D_j} - Gx_{M_j})^2 \quad +\alpha(Gy_{D_j} - Gy_{M_j})^2 \quad +\alpha(Gz_{D_j} - Gz_{M_j})^2 \end{aligned}$$

where α is a weighting factor. Although qualitative results are given it is not clear if a high gradient distance can reliably compensate for a low (wrong) gradient position distance.

The surface based representation of the model used in this work (opposed to the point wise representation in the above method) allows to take normal information into account and to restrict the search for corresponding points to the surface normal as shown in Figure 5.6 a. Two methods are investigated in order to detect a corresponding (data) shape boundary point lying on the model surface normal.

5.3.1 Correspondence based on threshold

As proposed in [Pic97], correspondence can be established by searching along the direction of the surface normal near the vertex, in both the positive and negative directions until a voxel is found that matches the bone threshold value. The exact location within a voxel is found by applying a quadratic line search root-finding [PFTV92] using trilinear interpolation. If a threshold is found in both the positive and negative search directions, both threshold positions are saved for further analysis, as described below.

In addition to matching the surface threshold value, the image gradient at the matching voxel must be consistent with the surface normal of the model. If the normals are not

consistent, it is quite likely that the surface that is being matched is invalid. For example, in Figure 5.7, the surface normal near the vertebral body surface at a given vertex may intersect the vertebral surface at three different points: two on the current vertebra (P1 and P2) and one on the adjacent vertebra (P3). It is assumed that in general only voxel (P1) will have an image gradient that is consistent with the vertex normal of the model.

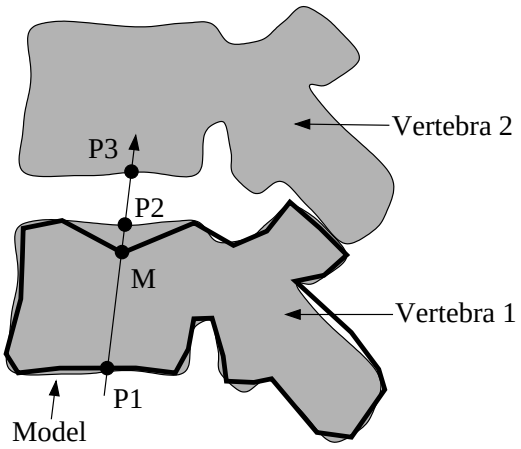


Figure 5.7: Because the model surface normal and the image gradient at the surface must be consistent, only $P1$ is chosen as the correct matching voxel. See text for details

The vertex normal of a vertex is defined as the average of the normals of all the triangles adjacent to that vertex. The above methods for isosurface extraction and triangle decimation ensure that the orientation of the normals is consistent for all vertices of the model.

Five distinct scenarios can occur when searching for correspondence between the vertices in the model and the image data. They are illustrated in Figure 5.6 b. The best-case scenario occurs when a vertex matches a nearby image point and the direction of the gradient in that point and the direction of the model vertex normal are consistent (P1). This point is considered to be a valid match. If the distance between the matched data point and the model vertex exceeds a (empirically) pre-defined limit Δ_{range} (P2, P3), then regardless of whether the surfaces are consistent or not, the vertex is considered to be a bad match. If the search distance falls within the search limit tolerance it depends on the angular deviation of the surface normal from the gradient direction if the vertex is considered to be a good match (P5) or a bad match (P4).

Unmatched vertices

After the initial surface matching is performed, a certain percentage of the vertices remain unmatched either due to an exceeded search range or due to an exceeded angular deviation. Instead of applying a local surface based relaxation technique to assign updated positions for these vertices as proposed in [Pic97], the new locations will implicitly be defined by the volumetric deformation applied to the entire model (section 5.4).

Hot Spots Regularization of the deformation ensures that few wrong correspondences only have a minor impact on the result. However, experiments show that in some cases wrong matches accumulate at the articulating facets of the vertebra (Figure 5.8, inferior and superior articular process) resulting in incorrect segmentation results. One possibility to deal with this problem is to infer the necessary deformation by the volumetric deformation guided by neighboring matched points. Those areas where this inferred deformation shall be applied can be defined interactively on the generic model.

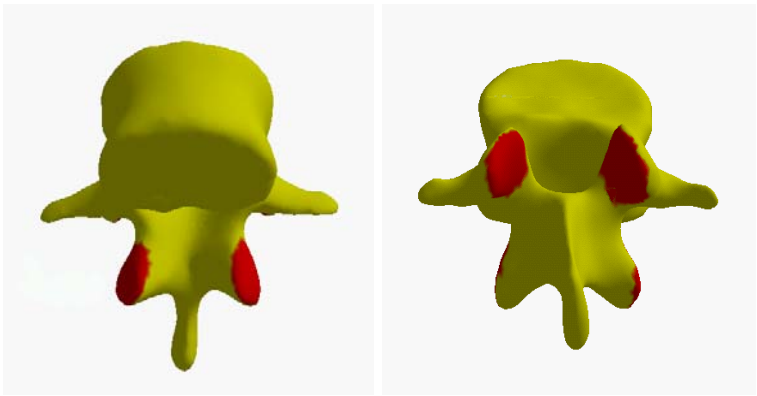


Figure 5.8: Hot Spots

5.3.2 Correspondence based on a statistical gray level model of the model boundary

The above method is only applicable for CT images, where a well defined bone threshold exists. However it is desirable to use the presented approach for automatic shape extraction also with other image modalities such as MRI.

Therefore a second more general approach for establishing correspondence between the model and the data image is investigated.

A set of pre-segmented training images is used to learn how the border of a specific model vertex should look like [ea99a]. This can be accomplished by sampling each image along the surface normal of the model matched to a set of training images around each vertex and subsequent construction of a statistical model of the grey-level distribution. It is important to note that by this procedure, no statistical information about the shape itself is gathered; it is only learned what a valid object boundary is.

Supposing that for a given point a profile of k pixels on both sides of the model point in the i^{th} training image is sampled, there are $2k+1$ samples yielding a vector $\mathbf{g_i}$. This is repeated for each training image, to get a set of samples $\{\mathbf{g_i}\}$ for the given model point. Under the assumption that they correspond to a single Gaussian distribution, their mean $\bar{\mathbf{g}}$ and covariance $\mathbf{C_g}$ can be estimated. This gives a statistical model for the grey-level profile about the point. This is repeated for every model point, giving one grey-level model for each model vertex. The quality of fit of a new sample, $\mathbf{g_s}$, to the model is then given by

$$f(\mathbf{g_s}) = (\mathbf{g_s} - \bar{\mathbf{g}})^{\mathbf{T}} \mathbf{C_g^{-1}} (\mathbf{g_s} - \bar{\mathbf{g}})$$

This is the Mahalanabonis distance of the sample from the model mean, and is linearly related to the log of the probability that $\mathbf{g_s}$ is drawn from the distribution. Minimizing $f(\mathbf{g_s})$ is equivalent to maximizing the probability that $\mathbf{g_s}$ is drawn by the distribution.

During search, a profile of m pixels on both sides of the current point ($m > k$) is sampled. Then the quality of fit of the corresponding grey-level model at each of the $2(m-k)+1$ possible positions along the sample (figure 5.9 b) is assessed and the one resulting in the best match (lowest value of $f(\mathbf{g_s})$) is chosen. This is repeated for every model point, giving correspondence for each model point.

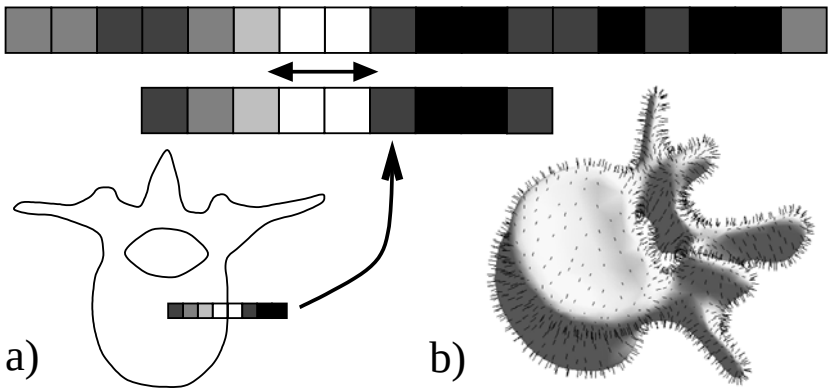


Figure 5.9: (a) The computation of a suggested movement for a single surface point. (b) Illustration of a vertebral shape with its outwards pointing profile vectors

5.4 Representation for the geometric transformation

The transformation $T_{\mathbf{p}}$ looked for, is the combination of a rigid-body transform R, and a multi resolution displacement function D:

$$T_{\mathbf{p}} = R \circ D$$

where $\mathbf{p}$ is a vector which gathers the 6 parameters that define R, and the deformation parameters that define D.

Assuming that $\mathcal{M} = \{M_i, i = 1...N_1\}$ and $\mathcal{P} = \{P_i, i = 1...N_1\}$ are the sets of matched model and data points, obtained by one of the two above explained methods the non-rigid registration algorithm minimizes the least-squares criterion defined in equation 5.1.

The deformation is modeled using a family of volumetric tensor product splines,

$$D = \sum_{j,k,l} \mathbf{u}_{jkl} S_j(x_i) S_k(y_i) S_l(z_i),$$

where the $\mathbf{u}_{jkl}$ are the spline deformation coefficients which comprise the parameter vector $\mathbf{p}$, and S_j, S_k, and S_l are spline basis functions. The $\mathbf{u}_{jkl}$ vectors are located on a regular 3D grid, where each basis function (a piecewise polynomial function) has a local support range. For the experiments reported in this dissertation a first order (trilinear) deformation spline is used. If the control points are left in their original positions, every point within the box will stay in its original position, too. By pulling a control point $\mathbf{P}_{ijk}$ to a new location $\mathbf{P}^*_{ijk}$, a smooth local transformation is imposed to the points within the box, as shown in figure 5.10.

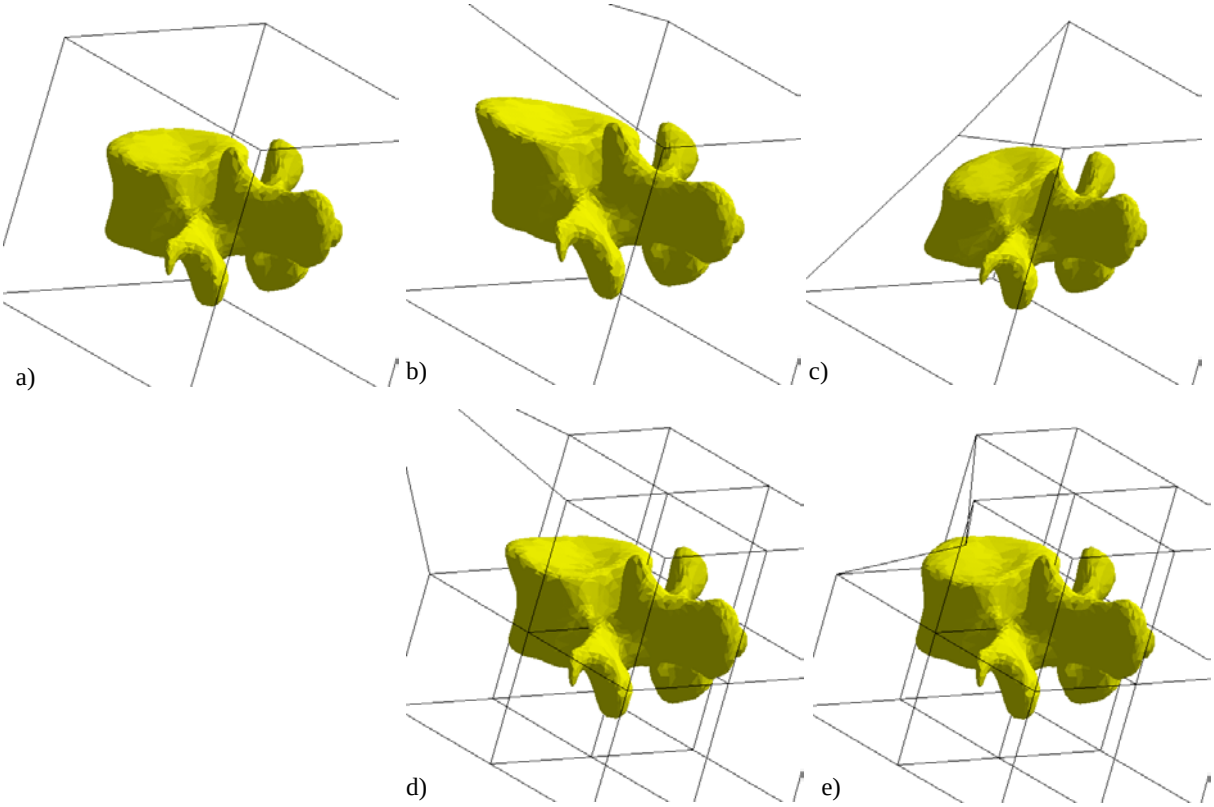


Figure 5.10: Hierarchical Free Form Deformations (FFD) applied to a vertebra model. Undeformed model (a), first level deformation (b,c), second level deformation (d,e)

The term $\mathcal{R}$ in equation 5.1 defines a regularization term which is applied to D in order to constrain the deformation and to obtain a smooth displacement function (see also section 4.2.2). As explained before the proposed method separates the representation of the model (triangle mesh) from the representation of the applied deformation (multi resolution FFDs). Thus regularization can be applied independently to either the deformation or the model itself. To regularize the deformation one can penalize for example the control point displacements from their original position, or the difference between displacements of adjacent control points. Regularization applied to the surface model can be performed by constraining the change of the model surface normal, the change of model surface curvature, or the change of the distances between model nodes (edge length) for

instance. A comparison study has been carried out in [ea99b]; authors conclude that good results can be obtained by applying any of the investigated methods. Furthermore they report that results are rather insensitive to the weight applied to the regularization term in equation 5.1. For experiments carried out in this thesis, the following regularization term that penalizes the control point displacement is used:

$$\mathcal{R}(\mathbf{p}) = \alpha \frac{1}{(a+1)(b+1)(c+1)} \sum_{i=0}^{a} \sum_{j=0}^{b} \sum_{k=0}^{c} dist(\mathbf{P}_{ijk}, \mathbf{P}^*_{ijk})$$

The regularization term is normalized to keep it independent of the number of grid points; α is a tunable weighting factor.

5.5 Optimization / Least square minimization

The optimization of $E(\mathbf{p})$ is performed using a conjugate gradient algorithm [PFTV92] and a multi resolution representation of T in order to smooth the solution and to speed up the minimization. Initial registration is performed by manual alignment. Subsequently the rigid-body transform parameters are estimated based on the concept of iterative corresponding point registration presented in section 4.4. Afterwards the nonrigid parameters are computed beginning on the coarsest level of the displacement grid. The subdivision is refined until a given resolution level is reached, (typically 32^3 or 64^3). Depending on the number of model points a few seconds or minutes are required to reach the convergence of the algorithm on an Alpha station.

When using a gradient descent technique such as the Conjugate Gradient method, there is a possibility that the minimization might fail because of local minima in the high-dimensional parameter space. By applying a coarse to fine strategy the parameters are always close to the optimal solution and thus the risk of local minima is reduced.

5.6 Results

To test the registration algorithm, experiments have been performed using a generic model of a lumbar vertebra and a set of lumbar spine CT images. Figure 5.11 shows the generic shape model (yellow points) and its corresponding points in a CT image for different values for Δ_{range} and Δ_{angle}. Valid correspondences are shown in green, invalid correspondences due to angular inconsistence are shown in red. The intersection of the model with the image plane is depicted as a thick red curve. Case (a) shows parameter values found to work well in practice, in (b) Δ_{range} is too low, in (c) Δ_{range} is too high, in (d) Δ_{angle} is too low, and in (e) Δ_{angle} is too high.

Figure 5.12 shows the generic shape model after manual alignment with a vertebra in a CT image (left), and after automatic rigid registration (right). One may observe, that the distances between corresponding points (the displacement vectors) decrease.

Figure 5.13 shows the effect of varying the bone threshold on the final registered model. In (a) the chosen threshold is too low, in (b) it is optimal and in (c) it is too high.

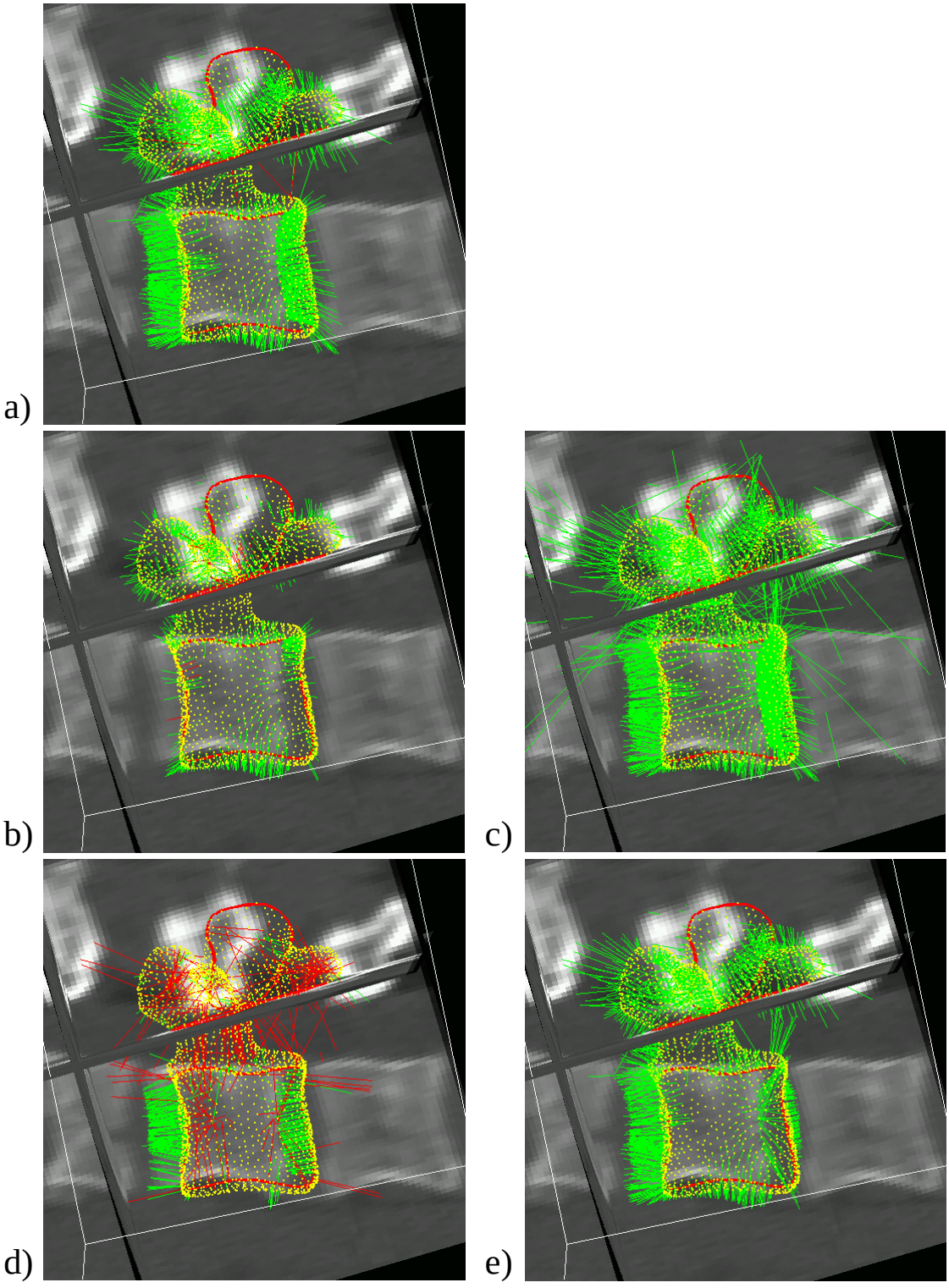


Figure 5.11: The generic shape model (yellow points) and its corresponding points in a CT image for different values for Δ_{range} and Δ_{angle}.(a) $\Delta_{range} = 15mm, \Delta_{angle} = 90°$, (b) $\Delta_{range} = 5mm, \Delta_{angle} = 90°$, c) $\Delta_{range} = 100mm, \Delta_{angle} = 90°$, d) $\Delta_{range} = 15mm, \Delta_{angle} = 15°$, e) $\Delta_{range} = 15mm, \Delta_{angle} = 180°$

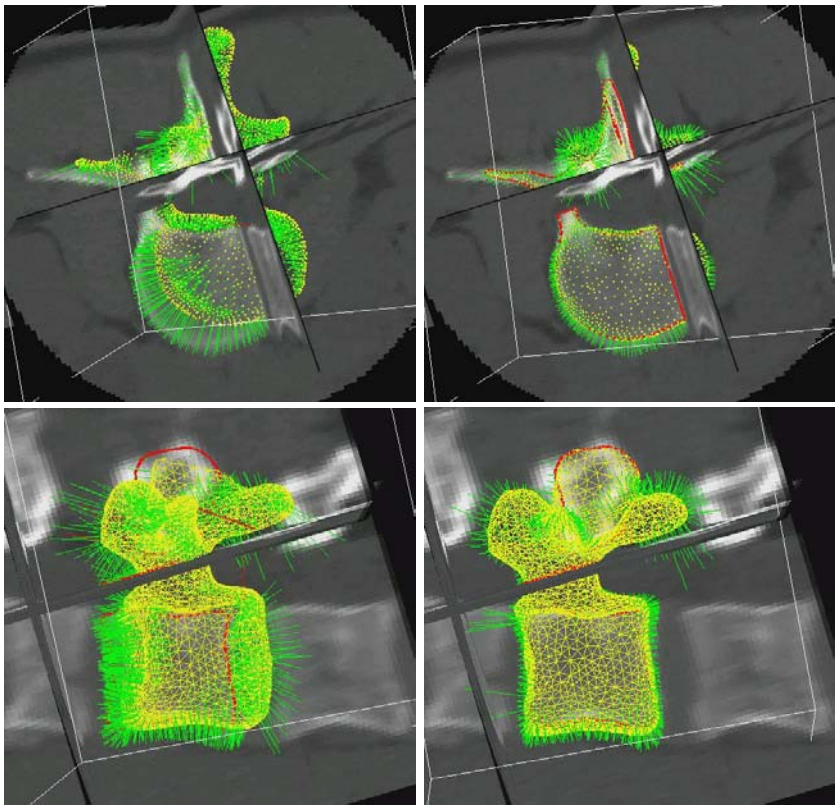
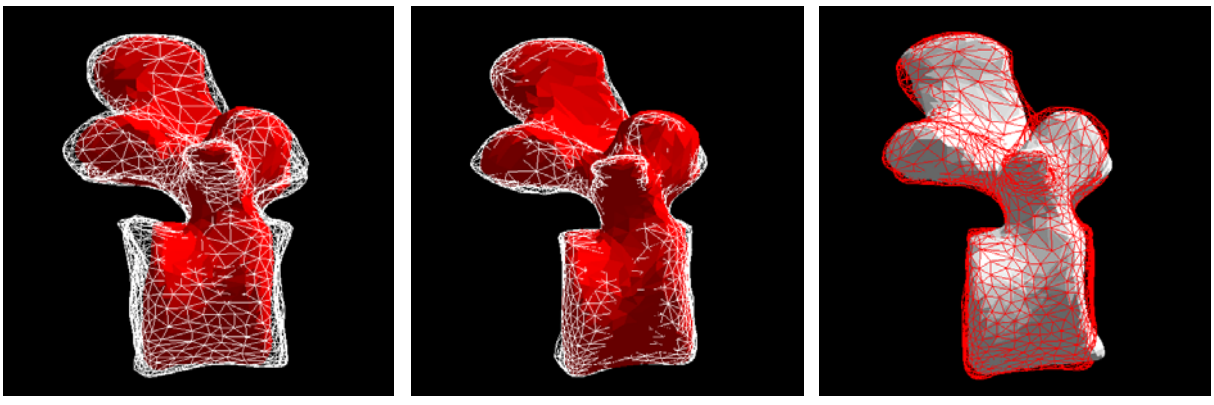


Figure 5.12: Rigid matching of the generic model with the CT exam (top axial view, bottom lateral view): after manual alignment (left), after automatic rigid registration (right)



a) b) c)

Figure 5.13: The influence of different bone thresholds on the resulting model. The mesh represents the registered model, the solid model represents the manually segmented CT model for comparison : (a) too low threshold , (b) optimal threshold , (c) too high threshold.

Figure 5.14 illustrates the evolution of the energy as a function of the number of iterations. The steps correspond to resolution changes of the deformation grid.

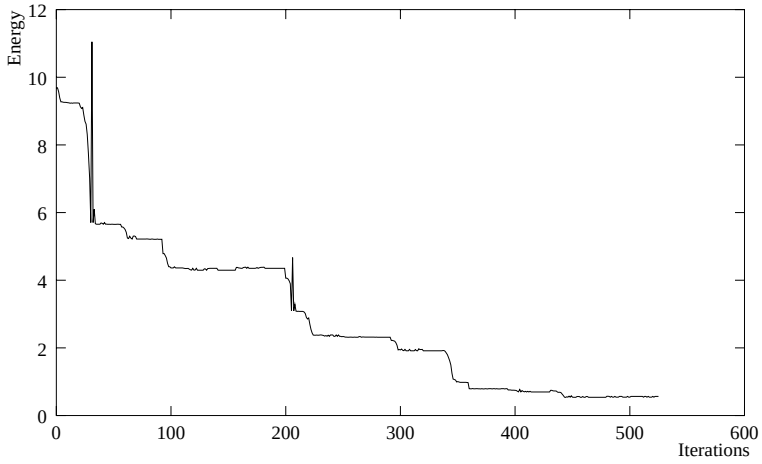


Figure 5.14: Evolution of the energy during function minimization

Figure 5.15 presents intermediate stages of the registration algorithm for different resolutions of the deformation grid. One may observe that the model boundary (red curve) gets closer to the vertebra for higher resolutions. One can also see the displacement of the FFD control points.

Table 5.1 summarizes RMS errors for 10 experiments. The average RMS error decreases from approximately $2.4mm$ after rigid registration to approximately $0.8mm$ after non rigid registration. Interactive prior selection of 'hot spots' as explained in section 4.4 yields slightly better results. Figure 5.16 visualizes the spatial distribution of the RMS error between the generic model and the underlying manually segmented CT-model after rigid registration (top) and after non-rigid matching (bottom). One can observe a considerably more homogeneously distributed error after non-rigid matching.

The overlap measure defined as the ratio between the volume of intersection and the volume of the union of the shape obtained by manual segmentation and the registered model, is very sensitive to even small differences in overlap, both inside and outside of the object model and is therefore a strong test for segmentation accuracy. It is calculated by comparing the voxelized generic model with the voxel map. Table 5.2 summarizes the results for the 10 performed experiments. The average overlap increases from approximately 64% after rigid registration to approximately 85% after non rigid registration. Slightly better results for interactive prior selection of 'hot spots' are confirmed by the overlap measure. Figure 5.17 visualizes the overlap for one case in three orthogonal cross sections. Figure 5.17 shows axial slices for an entire vertebra. First experiments performed with the statistical boundary model yield similar results as those carried out using the threshold based approach.

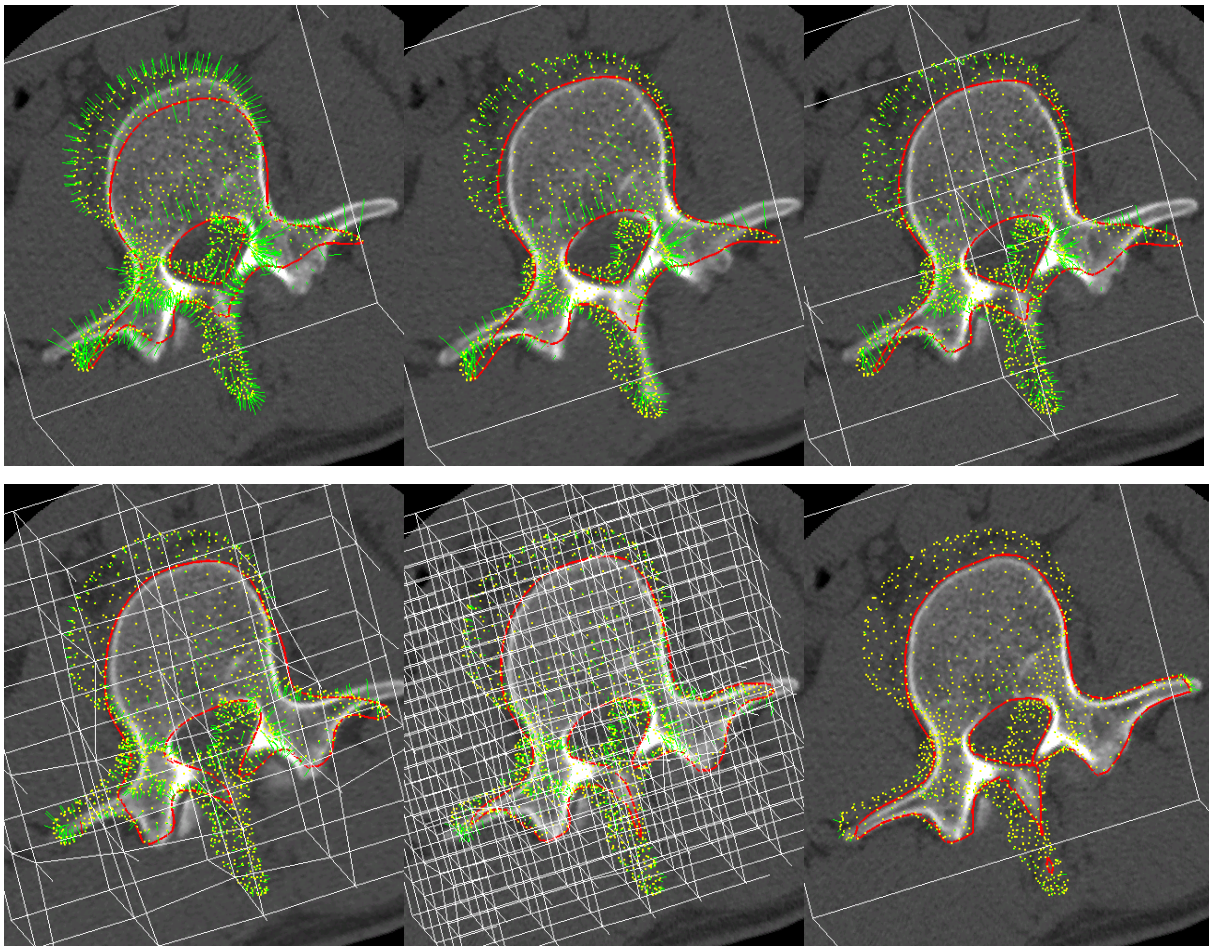


Figure 5.15: Intermediate stages of the registration method (from left to right, from top to bottom): after rigid registration, after first level deformation, after second level deformation, after third level deformation, after fourth level deformation, after fifth level deformation (shown without deformation)

Nr.	rigid RMS	non rigid RMS without hot spots	non rigid RMS with hot spots
1	1.46	0.64	0.61
2	1.74	0.72	0.69
3	3.26	0.79	0.74
4	4.16	0.98	0.93
5	2.42	0.61	0.65
6	1.67	0.66	0.63
7	1.58	0.88	0.85
8	3.29	0.63	0.63
9	1.37	0.89	0.77
10	3.38	0.86	0.75
mean	**2.43**	**0.77**	**0.73**

Table 5.1: RMS error for 3D segmentation (in mm)

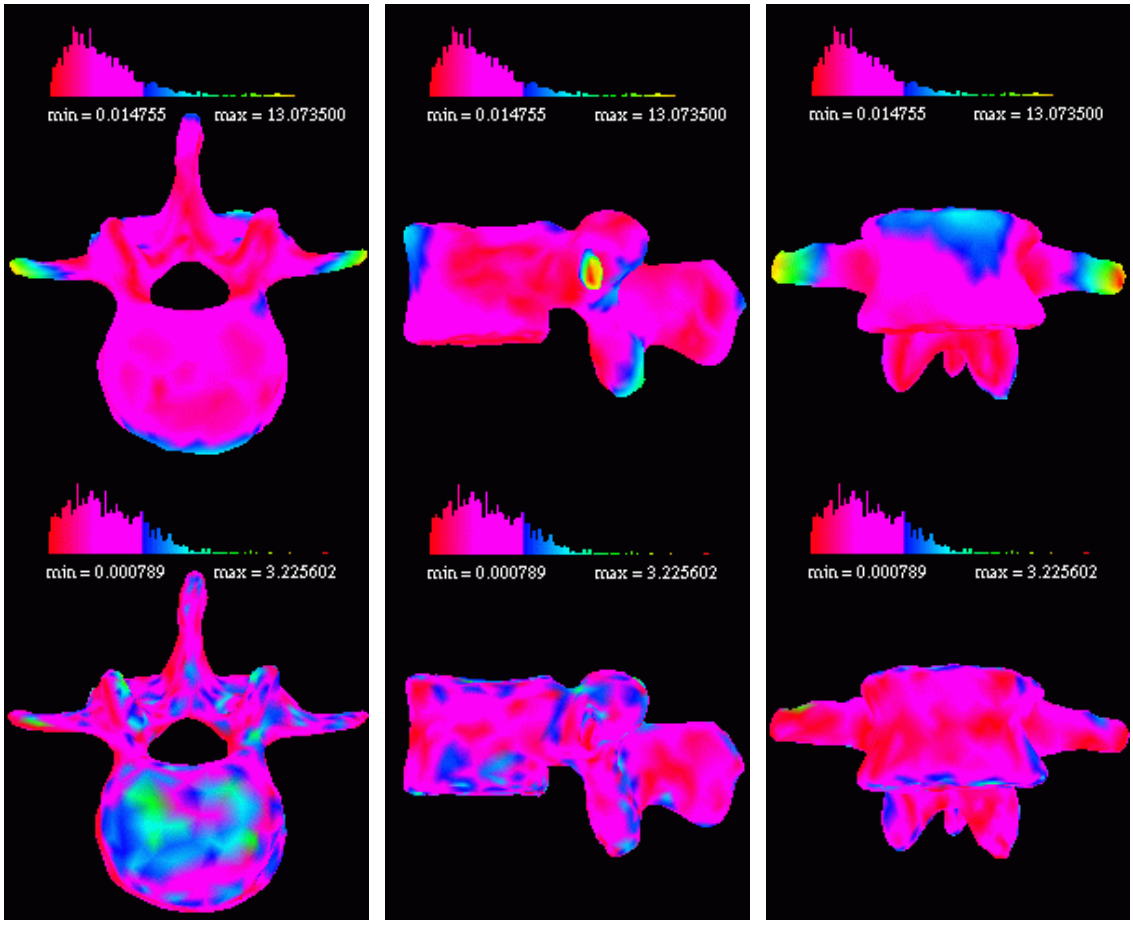
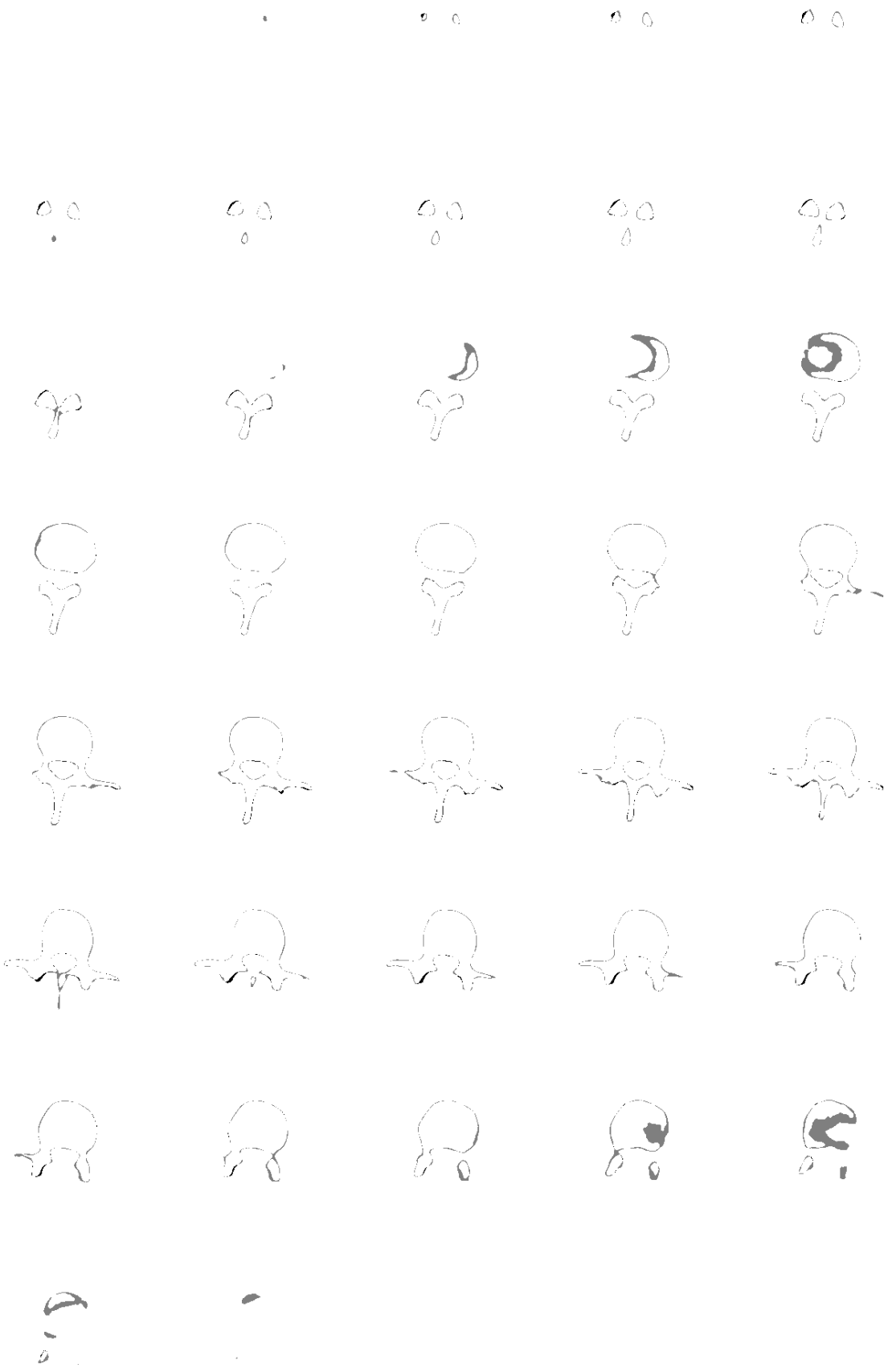


Figure 5.16: Error visualization for segmented vertebra axial, sagittal and coronal view: after rigid matching (top), after non-rigid matching (bottom)

Nr.	rigid overlap	non rigid overlap without hot spots	non rigid overlap with hot spots
1	61.69	86.34	85.80
2	59.32	85.69	86.04
3	66.21	86.82	86.94
4	57.74	84.53	85.78
5	62.29	85.52	84.54
6	56.67	83.82	84.92
7	58.01	79.43	80.36
8	65.05	89.29	89.10
9	60.08	79.35	81.25
10	63.97	84.67	87.12
mean	**61.10**	**84.56**	**85.19**

Table 5.2: Overlap measure for 3D segmentation(in percent)

Figure 5.17: The overlap measure can be computed by voxelizing the matched shape model and then comparing it to the manually segmented model. The figure shows an a) axial, b) sagittal and c) coronal slice through the resulting volume. Gray pixels are false positives and black pixels are false negatives.



5.7 Conclusion

In this chapter a method for automatic shape extraction from CT images has been presented. A volumetric coarse-to-fine registration method based on Free Form Deformations is used to match a generic model to the image data. Two methods methods have been investigated to attract the model boundary to the object contours in the gray-level image. The first methods is based on the bone threshold in the gray level image. Experiments performed on CT images of the vertebral column have shown that good segmentation results can be obtained except for regions where two neighboring vertebrae are articulating and thus the bone gaps are too narrow. The possibility to determine those areas prior to the registration manually and define the matching implicitly by neighboring mesh vertices and the volumetric deformation has also been studied and leads to slightly better results. A second approach based on a statistical gray-level model for each model vertex leads to similar results.

Chapter 6

Non-Rigid 3D/3D Registration of sparse scattered point data with a statistical shape model and its application to computer assisted ACL surgery

In this chapter a method for Non-Rigid 3D/3D Registration of intra-operatively acquired scattered point data with a statistically based shape model and its application to computer assisted reconstruction of the anterior cruciate ligament is investigated. The rupture of the ACL has become one of the most common knee injuries. One problem during reconstruction is to find the optimal attachment points for the graft. Therefore a system for computer assisted reconstruction of the ACL has been proposed at TIMC laboratory [JLD98]. During surgery the surgeon collects several data points on the tibial and femoral joint surface with a 3D localizer system. These 3D data are used to find those attachment points resulting in a low anisometry of the graft while preventing impingement between the graft and the femoral notch. As the collected data points only cover a small surface patch of the femur, it is desirable to extrapolate these data to have also a visualization in those areas where no data points are available. A sufficiently good approximation of the actual femur by a model would further allow to better deal with the notch impingement problem of the graft. The chosen approach is to fit a deformable model to the data points. It can be subdivided in 2 steps, constructing the model and fitting this model to the data. To incorporate a priori knowledge into the model, the allowed deformations are determined by the statistics of the shape variation of a set of training objects. Matching the training objects together is obtained by elastic registration of surface points using octree-splines. The fitting process of the sparse intra-operative data with the statistical model results in a nonlinear multidimensional function minimization. Experimental results with a model generated from 10 dry femurs are presented, including fitting of the model with both simulated and real intra operative data.

The rest of this chapter is organized as follows: In section 6.1 the current status of the

system that has been developed for computer assisted ACL reconstruction is detailed and the motivation for the work of statistically based shape modeling that is presented in the rest of the chapter is further clarified.

Section 6.2 provides a brief overview of the work in the domain of surface reconstruction using scattered point data. In section 6.3, an original method to build a statistical surface model using a population of samples represented by unorganized sets of 3D points is proposed. Then the fitting method of the model to the point data is presented in section 6.4. Results are shown in section 6.5 and discussed in section 6.6.

6.1 ACL reconstruction

6.1.1 Background

The rupture of the Anterior Cruciate Ligament (ACL) has become one of the most common injuries among young athletes. Approximately 50,000 reconstructions of this ligament are performed each year in the United States [FJ97]. Because of the instability of an ACL deficient knee and the risk of secondary damage, an increasing number of them need to have reconstructive surgery. During surgery a substitute ligament is implanted between the femur and the tibia to restore stability. Most frequently, ligamentous tissue of the patient is used for the graft, such as the middle third of the patella tendon or the hamstring tendon of the injured knee. The graft is inserted and fixed into a femoral and tibial tunnel according to two main criteria, low anisometry of the graft and avoiding notch impingement (see section 2.2).

The following criteria also influence the short and long term results of ACL reconstruction but are not further addressed here.

- Choice of the graft material (autologous, homologous, synthetic) [Des96]
- Tibia and femur tunnel orientation [SdGD97]
- Tunnel length
- Conflicts with other intra-articular structures[Des96]

6.1.2 Computer assisted technique for ACL reconstruction

A system has been investigated which allows positioning of the central part of the ligament graft at the least anisometric sites, while preventing notch impingement [DLO+95, DLJ+95]. It provides the surgeon with the predicted anisometry and the profile of the length variation of the graft as a function of flexion angles for any attachment points on the femoral notch and tibial surfaces. A demonstration of the clinical interest of this method is presented in a study performed on 23 patients [JLD98].

The system uses only *intra operative* data obtained with an optical localizer (Optotrak, Northern Digital, Toronto). This system enables the surgeon to digitize 3D points interactively, to track relative bone motion, and to locate the pose of surgical tools in real

time. Dynamic reference rigid bodies made of infra-red LEDs are attached to the bones, the pointers, and the tools (Fig.6.1). Neither pre-operative CT or MRI exams, nor pre- or intra-operative X-ray images are required. The method can be divided into three steps:

Passive Flexion-Extension

At the beginning of the surgical procedure, the relative movement of the femur with respect to the tibia is captured with the optical localizer system by applying a passive flexion-extension. A series of matrices $M_{\theta i}$ describing the transformation between the femoral reference system Ref_F and the tibial reference system Ref_T is computed for about $N = 20$ flexion angles of the knee $\theta_i, i = 1...N$.

3D Points Acquisition

The surgeon interactively collects surface points arthroscopically on the femoral and tibial joint surface using a 3D optical pointer (Fig.6.1). In practice, 3 areas are digitized on the flexed knee:

area A: Femoral surface candidate for insertion

The surgeon acquires $N_f = 20$ to 50 surface points $f_i, i = 1...N_f$ on the femoral notch, in an area that corresponds to all the possible candidate points for the femoral attachment site. These points are approximated by a bicubic spline patch s_f, in the form of $z = s_f(x, y)$ where (x, y) denotes the least-square fitting plane of the points f_i.

area B: Anterior border of the femoral notch

The surgeon digitizes a few points on the anterior part of the notch, which corresponds to a horse shoe shaped 3D curve that defines the location of possible conflict (impingement) between the graft and the femoral surface. The digitized points are linearly interpolated to provide a 3D curve. Each line segment defines the axis of a cylinder with a 3-mm radius, which is represented graphically as a thick curve in order to visualize possible notch impingement with a clearance of 3 mm, as recommended by some authors [YDP92].

area C: Tibial surface candidate for insertion

The surgeon acquires $N_t = 20$ to 50 surface points $t_i, i = 1...N_t$ on the tibial plateau, in an area that corresponds to all the possible candidate points for the tibial attachment site. These points are approximated by a bicubic spline patch s_t, in the form of $z = s_t(x, y)$ where (x, y) denotes the least-square fitting plane of the points t_i.

The approximate accessible area on the femur during arthroscopic surgery only includes the condyles and the femoral notch (Fig. 6.2).

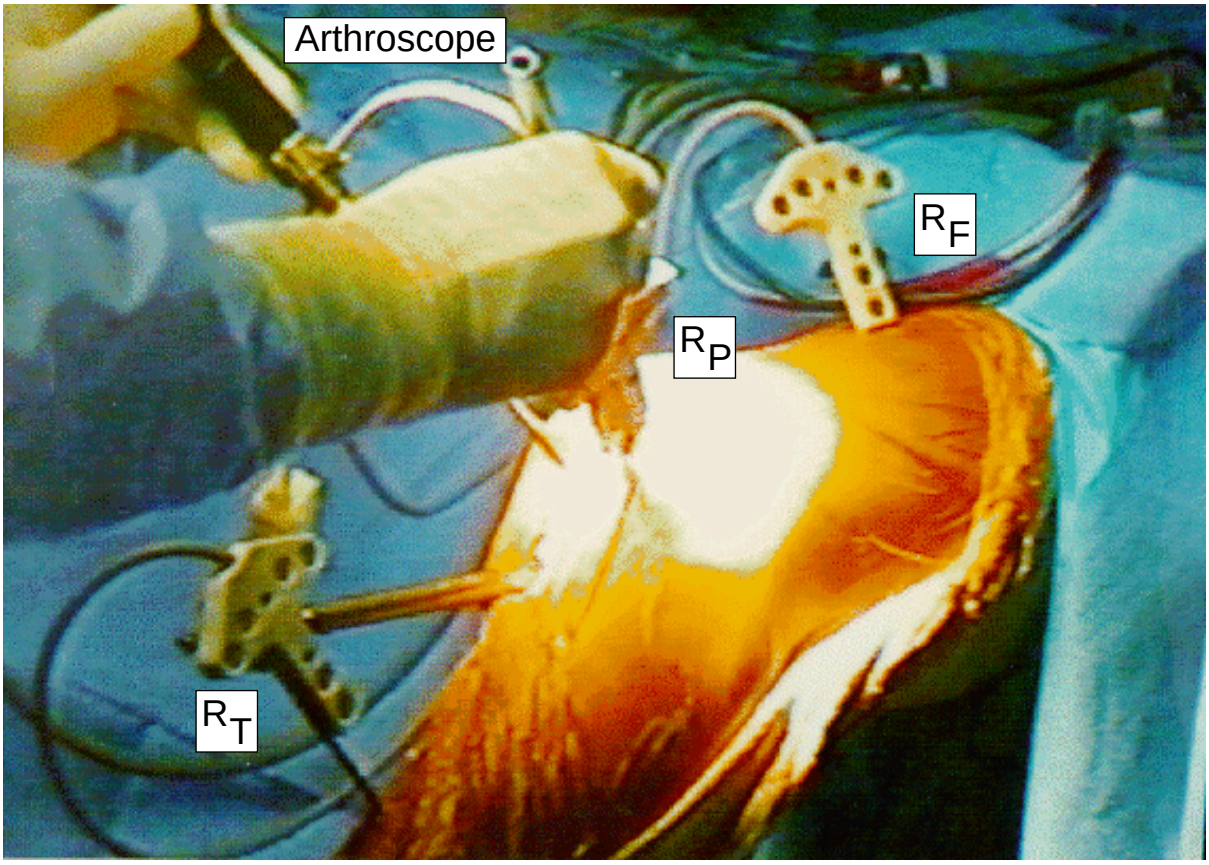


Figure 6.1: Surface digitizing during surgery. R_F and R_T denote rigid bodies attached to the femur and tibia using pin fixation in the bones. R_P denotes a rigid body attached to a pointer for surface digitization. Several infra-red LEDs are attached to each rigid body. Thus the position and orientation of R_F, R_T and R_P are tracked in real-time using an external optical localizer.

Interactive selection of the attachment points

Based on the geometric and kinematic data acquired with the optical localizer, the system displays 3D views and images that enable the surgeon to optimize anisometry and notch impingement criteria in real-time based on the following information provided by the computer:

- **Anisometry profile**

  For any pair of points F and T on the interpolated femoral and tibial spline surfaces, the system can compute the predicted ligament length variation curve. The tibial point is defined in Ref_F for any flexion angle $\theta_i, i = 1...N$ by $T_i = M_{\theta_i} T$. Therefore, the length of the graft fiber at the flexion angle θ_i is $d_i = ||FT_i||$, and the variation of length with θ_i can be displayed; it is referred to as the anisometry profile (Fig.6.4). Anisometry is defined as the variation of length along flexion-extension:

$$ANI(F,T) = \max_{i=1..N} d_i - \min_{i=1..N} d_i \tag{6.1}$$

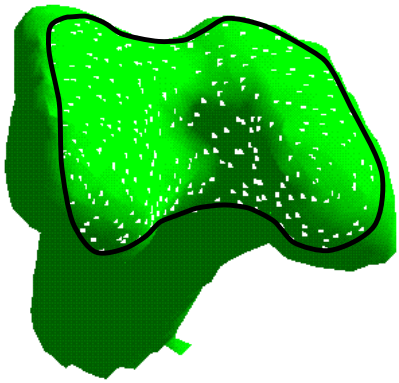


Figure 6.2: Maximal accessible area for data acquisition on the femoral surface: distal part of the condyles and intra-articular notch.

- **Anisometry maps**

  For a given point T_0 on the tibial surface, ANI is a scalar function of F. Therefore, it can be represented as a pseudo-color map that plots the values of $ANI(F, T_0)$ on the femoral surface. This defines the femoral anisometry map. A value of $ANI(F, T_0)$ below 2 mm is represented as dark green, a value between 2 mm and 4 mm is represented as light green, etc. (Fig.6.3). Similarly, for a given point F_0, the system computes the tibial anisometry map $ANI(F_0, T)$. Both femoral and tibial anisometry maps can be projected on the (x,y) planes that represent the definition domains of the spline patches. This defines the developed anisometry maps (Fig.6.4).

- **Graft envelope**

  For any pair of points F and T, the computer predicts the envelope of the graft as follows: In the surgical technique considered in this paper, the surgeon drills the tibial tunnel from the outside with a knee positioned at 90° of flexion, then the surgeon drills the femoral tunnel from the inside by passing a drill guide through the tibial tunnel. Therefore, the system selects the matrix $M_{\theta=90°}$ which reflects the closest position to 90° of flexion. For this position of the knee, a cylindrical tunnel centered in points F and $M_{\theta=90°}T$ with 5-mm radius is drilled virtually in the femoral and tibial surfaces, i.e. that the intersection curves between the infinite cylinder and the surfaces are generated (Fig.6.3). Considering the line segments that link the intersection points on the femoral and tibial surfaces respectively, one obtains a representation of the external envelope of the graft that can be displayed for any flexion angle of the knee. In particular, the envelope generated for this 90° position is flattened when the knee is in extension, which reflects the real behavior of the graft. Note that this construction makes it possible to consider various fibers of the graft in its periphery instead of a single central graft. By placing the virtual model in full extension, the surgeon can observe in lateral and axial 3D views if there is a risk of intersection (impingement) between the graft envelope and the

anterior border of the femoral notch (Fig.6.3).

The system is initialized with two points F and T selected in the middle of the femoral and tibial spline patches and the surgeon can navigate interactively in both surfaces using a pointer or any surgical tool such as a drill equipped with a rigid body that is localized in Ref_F or Ref_T. First, by pressing a foot switch, a tibial point T_0 is selected such that the 3D view does not show lateral or axial impingement between the graft envelope and the anterior border of the notch. Then, the surgeon navigates on the femoral surface until a point F_0 is found with a low anisometry and a profile that corresponds to loosening along with knee flexion. This process is iterated if necessary.

6.1.3 Definition of the problem

Fig.6.4 shows the graphical user interface of the current system for computer assisted ACL surgery with the anisometry maps, and a 3D view with the spline surfaces. As only a small surface area has been digitized and no pre-operative data such as CT or MRI are used, it is very difficult to recognize the actual pose of the tibia and the femur only from the small surface patches generated by the bicubic splines. In order to provide the surgeon with a more complete and realistic view of the scene it is desirable to have a visualization of the whole femur (respectively tibia) as shown in fig. 6.5 (only the femur is visualized).

As illustrated in Fig. 6.6, the addressed problem is to recover the complete shape of the bones (proximate to the joint) from the few available data points.

More precisely the objectives are quadruple:

- Visualization of a complete anatomical model

  As already mentioned, the primary objective is to provide the surgeon with a better visualization of the complete relevant anatomy, including not only the shape of the femur and tibia, but also any information contained in the model such as anatomical landmarks for instance.

- Post-operative referencing

  Fitting a complete model with a few data points also means that the results of each surgical case can be reported in the model post-operatively, which is an essential element of clinical studies.

- Shape-based interpolation

  Performing shape-based interpolation such that the least a minimal number of points has to be digitized manually by the surgeon. This aims at saving time which is critical for this surgical procedure. For example, the delineation of the anterior part of the notch (area B) is currently performed manually but it is rather delicate because the edge is not sharp. Using the proposed technique, it becomes possible to consider the complete surface in the area of the anterior part of the notch, and not only an approximate curve, as candidate for possible impingement.

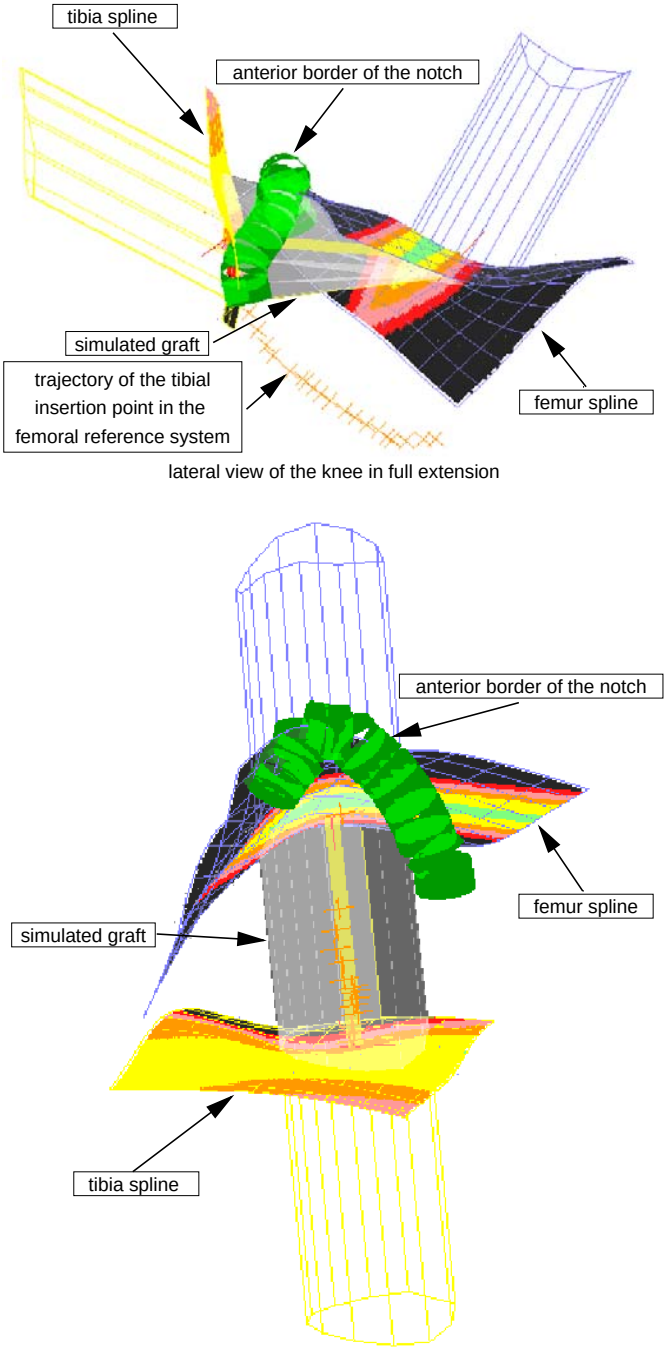


Figure 6.3: Lateral (up) and axial (bottom) views of the 3D objects reconstructed by the computer from raw data. Those images show the femoral and tibial spline patches on which anisometry maps have been superimposed. In this example, the envelope of the graft displayed in a transparent mode does not intersect the anterior arch of the femoral notch (horse shoe shape).

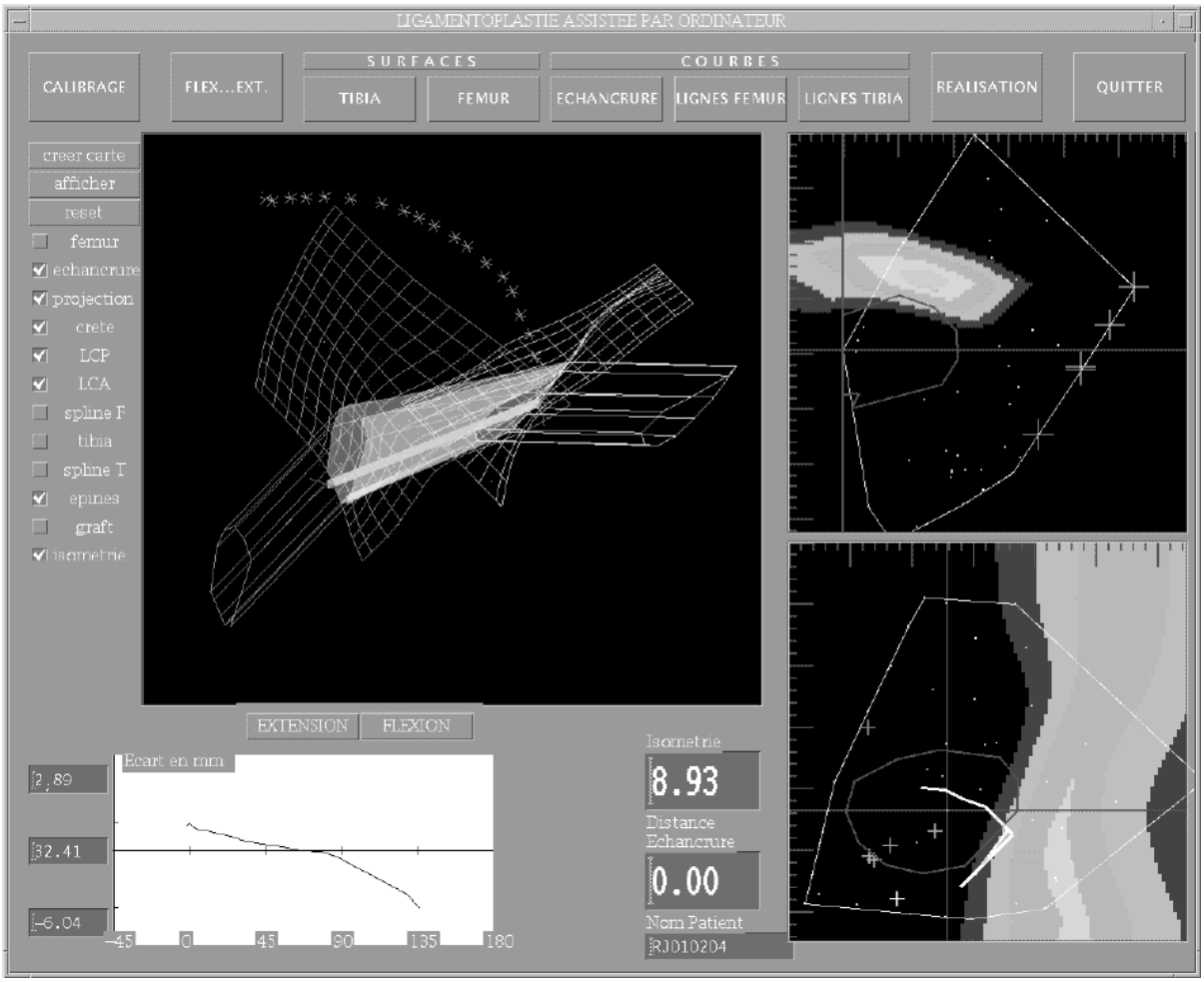


Figure 6.4: Graphical User Interface (GUI) for computer assisted ACL reconstruction. Top left: 3D view including femoral and tibial surfaces, with the predicted graft envelope. Bottom left: anisometry profile. Top right: developed femoral anisometry map. Bottom right: developed tibial anisometry map.

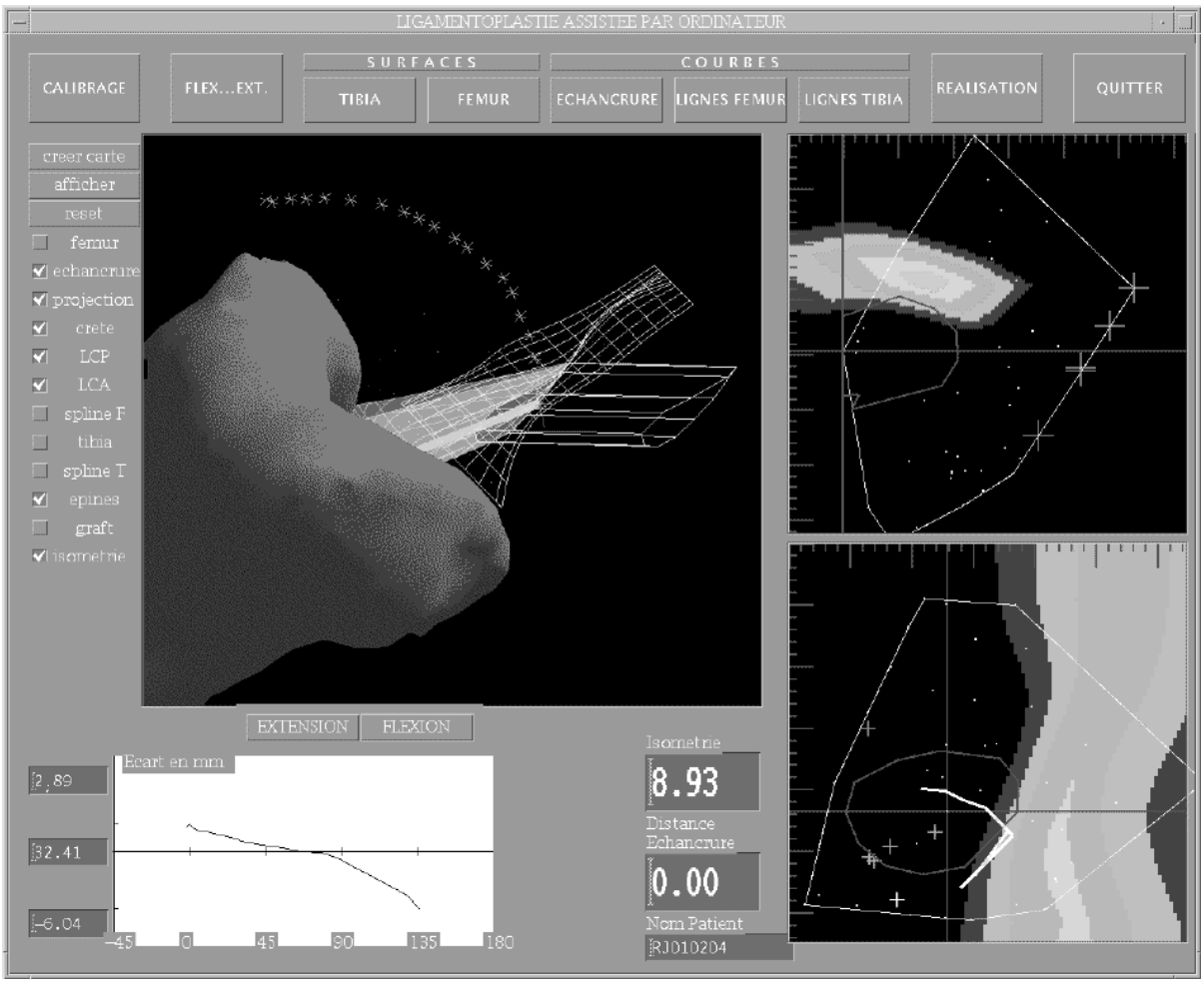


Figure 6.5: GUI using a complete model of the femur instead of a bicubic spline in the 3D view window (up left). The global orientation and anatomy of the knee can be understood.

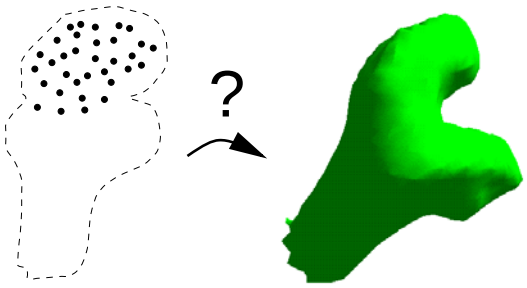


Figure 6.6: Problem: How to extrapolate a few data points to a complete femur?

- Extrapolation

  Assuming the resulting model is of sufficient accuracy outside the digitized areas, the extrapolated shape can be used to define additional criteria for optimal ACL reconstruction. For example, using a complete surface of the bone, the length of the tunnels in the bones can be easily determined, which is important to predict and compare with the graft length before drilling. If necessary, additional points can be acquired in the external parts of the bones using transcutaneous direct or ultrasound-based digitization for instance.

In this thesis experiments are conducted concentrating on the isolated femur, but the approach can be applied to the tibia as well, or preferably to the pair of femur and tibia in extension.

6.2 Related Work

The literature of computer vision proposes a large variety of methods for shape reconstruction from scattered object data, also referrered to as scattered data interpolation. An overview is presented in [BV91]. More recent references can be found in [OG98, GM97].

Here, the problem is to fit a surface to scattered, unorganized data associated with an isolated object. The data is unorganized in the sense that the adjacency relation between surface points is not known. When surface data is obtained from a laser surface scanner for instance, the relationship between surface points is usually known by virtue of the methodical way in which an object is scanned. Consequently, many techniques for recovering surfaces from range data require the adjacency relationship to be known.

Following [MM97], methods for surface reconstruction from scattered point data not relying on this information can be based on spatial subdivision, on distance functions, on warping and on incremental surface growing. One example for a surface oriented subdivision technique is the work by Algorri and Schmitt [AS96]. In a first step of their algorithm a rectangular bounding box of the given data set is subdivided by a regular voxel grid. In a second step, the algorithm extracts those voxels which are occupied by at least one point of the sampling set. Subsequently, the outer quadrilaterals of the selected voxels are taken as a first approximation of the surface. Afterwards the surface can be transferred into a triangular mesh by diagonally splitting the quadrilaterals. The work in [Boi] belongs to the volume-based subdivision approaches, and relies on the removal of tetrahedra from the tetrahedral decomposition of the objects convex hull. 3D alpha shapes used in [EM94] are a kind of generalization of the latter method. In order to reconstruct roughly an object a user must provide an appropriate value of the parameter alpha. The method by Hoppe [HDDW92] is based on the evaluation of a signed distance function f which is defined for each point in the space such that f estimates the signed geometric distance to the unknown surface. This is accomplished by associating an oriented plane with each of the data points, which serve as local linear approximations to the surface. The function has positive, negative, or zero values for points outside, inside, or on the border of an object, respectively. The zero set $Z(f)$ is the estimate of the surface. In the second stage a variation of the marching cube algorithm is used to extract the isosurface

$Z(f)$ from the scalar function f. See [MM97] for an overview. However, none of these references can be applied to the here present problem , because a) they deal with surfaces for which the global shape is usually unconstrained, b) they usually require the point set to be sufficiently dense and to have a more or less uniform density, and c) they are not capable of accomplishing shape extrapolation.

In particular because of the latter reason literature relevant to deformable models is more appropriate to the present problem. An introduction to deformable models was given in section 4.3. Most of the methods, which are based on the snakes formalism [KWT88], are successful for interpolating data but they cannot perform extrapolation in a reliable manner. Hence models capable of incorporating more specific priori knowledge have to be considered. One class of methods use parametric models based on deformable templates [YHC92]. Such templates may be built from sets of primitives such as circles, lines or arcs each of which has some degree of freedom to move relative to the others. However, this method is *not general* - it is difficult to apply using general a priori knowledge, i.e. a new template and fitting scheme must be produced for each application. Another possibility is to consider models such as deformable superquadrics [MT93]; however those models are appropriate to capture shapes defined by many data but not to extrapolate local data (the superquadrics convey information about the global shape but this part of the model is not accurate enough for the present application). Similarly, using volumetric deformations with regularization constraints such as presented in [SL96] can be expected to preserve the shape of an organ, but this will be true only in the local neighborhood of the collected data. Statistical shape models permit to provide a "reasonable" shape estimate not only in those areas covered by the acquired points but also outside. One possibility is to use statistical models based on Fourier representations, such as [SD92, SKBG96]. Another method is based on a decomposition of shapes into a basis of fundamental deformations using modal analysis [PS91]. It is also possible to use features such as crest-lines and to perform modal analysis on the features [STA96].

In section 4.3.3 a statistical shape model with modal representation based on principal component analysis directly applied to the nodal representation of the shape contour [CTCG95a] has been presented and will be used for the present problem.

6.3 Building a 3D statistical shape model from point data

Building the statistical model requires 3 steps:

- Acquiring the training shapes
- Establishing point to point correspondence between all training shapes
- Principal component analysis

These three steps are detailed in the following.

Figure 6.7: Shape database comprising 11 right femurs

6.3.1 Acquisition of Training Shapes

In chapter 5 a highly automatic method for extracting training shapes from CT-images including automatic landmark placement was presented. For the first experiments addressing the issues related to the computer assisted ACL-surgery system presented in this chapter, a small population of 11 dry right human femurs was available, see Fig. 6.7. Ten of them had been digitized manually using a 3D optical localizer Optotrak, resulting in 10 point sets, each representing the surface of one femur as a non-organized set of points. Each point set contains approximately 1500 points randomly distributed on the bone surface (while covering approximately the same distal femur area) thus constituting a point set $\mathcal{P}^j = \{P_i^j, i = 1...N^j\}$ for each case $j = 1..10$.

The available additional eleventh femur was considered separately as a template. A larger area is digitized in this case as it must be guaranteed for the matching algorithm presented in the next section that the 10 training shapes are a subset of this template. The algorithm proposed by Hoppe [HDDW92] (see section 6.2) is used to compute a triangle mesh from this point cloud. Since it requires a high point density, the template was digitized with a higher density of approximately 6000 points.

Subsequently the resulting triangle mesh is decimated and smoothed as explained in section 5.2.4. The result is a triangular mesh TM_{ref} of approximatively 1500 vertices.

6.3.2 Definition of a point to point correspondence between Training Shapes

Now each of the 10 point sets $\mathcal{P}^j$ has to be matched to the template mesh in such a way that each vertex of the template mesh TM_{ref} is mapped to its anatomically corresponding point on the femur represented by the set of points $\mathcal{P}^j$.

This alignment and matching process is accomplished using a multiresolution approach proposed by Szeliski and Lavallée, based on hierarchical adaptive space subdivision tech-

nique called octree-splines [SL96]. The method performs a least squares minimization of the distances between a sparse and unorganized set of points and a dense set of points used to build a 3D octree-spline distance map [LSB91]. In the present case, the dense set of points is obtained by resampling each facet of the template triangular mesh TM_{ref} up to 10,000 points.

The deformations between the 3D surfaces are described as a warping of the space containing one of the surfaces based on concepts from free-form deformations, octree splines and hierarchical basis functions. This technique does not require the extraction (manual or automatic) of features on the two surfaces. Fig. 6.8 shows the octree containing the two shapes to be registered after rigid registration (top) and after non rigid registration (bottom).

The result of the octree-spline based registration technique is a smooth volumetric transformation T that maps every point P_i of the actual data space to a point $M = T(P_i)$ of the model space. As T is isomorphic, it is possible to inverse this transformation. Given a point M of the template mesh, an iterative search is performed to find the point $P_i' = T^{-1}(M)$ such that $||M - T(P_i')||$ becomes smaller than a given bound ϵ. By this process, each point M of the template mesh is assigned to a data point P_i' for each data set j of the N data spaces. Note that the points P_i' were not in the data sets initially, but they were implicitly interpolated using the octree-spline deformation. The iterative search is necessary, because there is no analytical expression for T^{-1}. For the experiments ϵ was chosen to be $0.01mm$.

In the case where the data points would have been represented by dense sets of points (with a higher density than the template model), one could have applied the octree-spline mechanism directly from the model towards the data. But the presented technique is more general in the sense that it only needs one single case (the template) with guaranteed high point density but allows holes or non dense regions in the rest of the population, possibly collected by different means (direct digitization, image segmentation, etc.).

The *mean shape* which is necessary for further computation is computed using an iterative algorithm. At first all training shapes are matched to the template mesh as described above. After calculating the mean shape, now all training shapes are matched to this current mean. This process is repeated until convergence occurs. This approach avoids proceeding with a mean calculation based on matching results between shapes possibly differing too much in shape, thus causing incorrect point correspondence within the matching algorithm. Matching each training shape only with the mean rather than with another training shape helps avoiding this problem as shape differs less between the mean and a training shape than it may differ between two arbitrary training shapes. This issue has been addressed in further detail in [ea01a] for instance. However, using the given population, convergence occurs after 2-3 iterations. Subsequently the deformation modes can be computed by applying PCA to the data as epxlained in section 4.3.3.

6.4 Model Fitting

To recover the whole surface of an object given few sparse data, it is necessary to find the rigid transformation (rotation $\mathbf{R}$, translation $\mathbf{T}$) between the data and the model and the

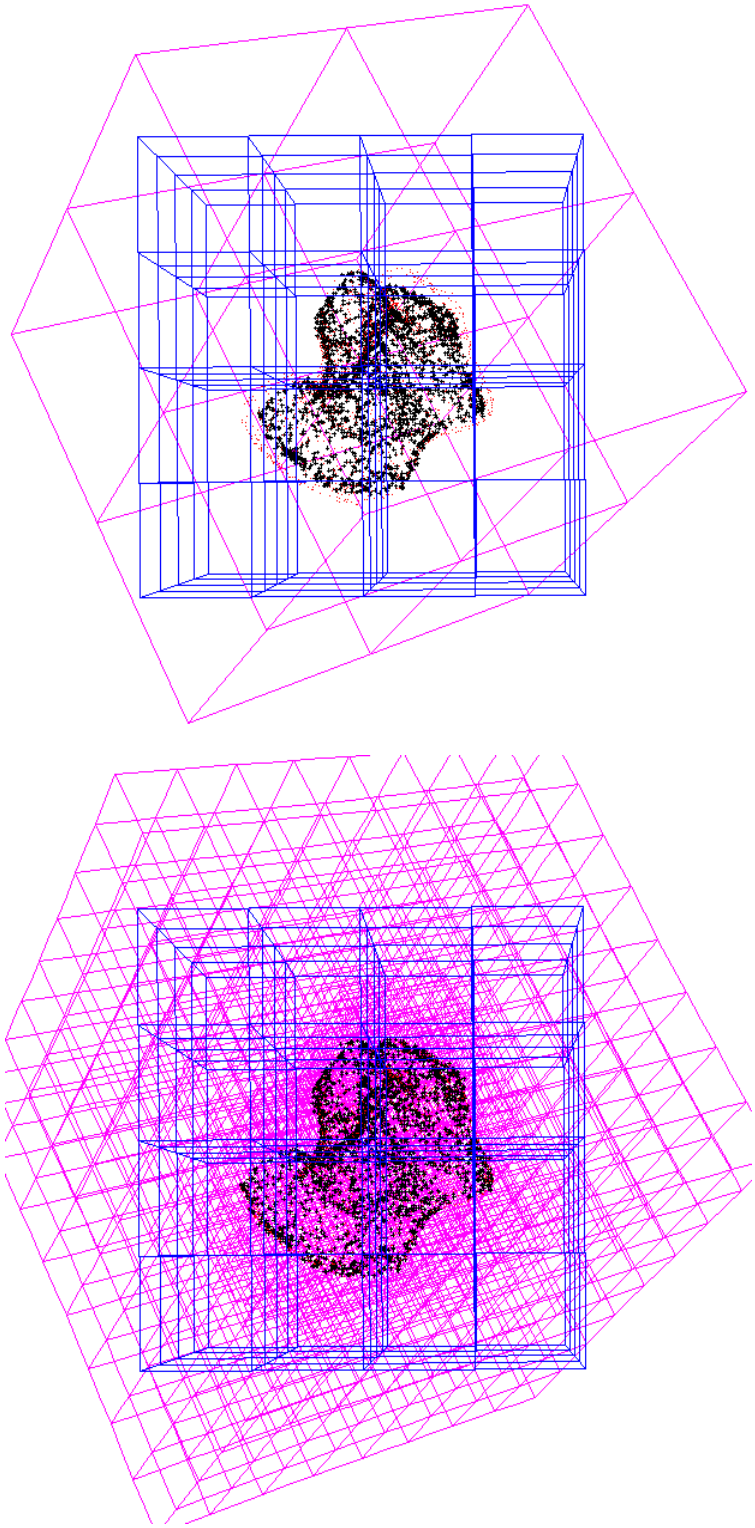


Figure 6.8: Hierarchical volumetric deformation of a low density training set with a high density template mesh, using octree splines. Top: after rigid alignment. Bottom: after elastic registration.

decomposition of the t preserved eigenvectors in such a way that the distances between the data points and the model are minimized. The objective function to be minimized is defined as follows:

$$min \quad f = \sum_{i=1}^{D} min(\|\mathbf{d}_i - \mathbf{m}_j\|) \quad _{1 \leq j \leq M} \tag{6.2}$$

with $\mathbf{m}_j = \mathbf{R}(\bar{\mathbf{m}} + \sum_{i=1}^{t} w_i \mathbf{e}_i) + \mathbf{T}$, D the number of data points, and $\mathbf{d}_i$ the i^{th} data point. The computationally most expensive step in the registration process is finding the closest point in the model to each data point. The computational complexity evaluating f is $\mathbf{O}(DM)$ using exhaustive search. It is not possible to speed up the computation by pre-computing a 3D distance map as proposed in [SL96] for instance, because the model to be registered is deformed at each iteration.

Therefore k-dimensional binary trees (3-d trees, in this case k=3) are used to speed up the computation [FBF77].

The k-d tree is a binary tree in which each node represents a subset of the data records (3D points in the present case) and a partitioning of that subset. Each nonterminal node has two children that represent the two subsets defined by the partitioning. The terminal nodes represent mutually exclusive small subsets of the records. A 3-d tree divides space into a collection of rectangular parallelepipeds that correspond to the terminal nodes. This data structure provides an efficient method for examining only those points closest to a given point. Fig. 6.9 shows exemplarily the space partition by a two dimensional k-d tree.

A k-d tree can be constructed in $\mathbf{O}(M \log(M))$. Searching the closest point in the tree to the given data points can then be performed in $\mathbf{O}(D \log(M))$. Notice that for each function evaluation the k-d tree has to be reconstructed since using different weights w_i for the shape parameters results in a different point distribution.

As the model is a triangular mesh surface, computation of the closest point on the model surface requires an additional step. The output of the k-d tree based search will return the vertex of the triangle mesh which is closest to the data point. Assuming that the closest surface point lies within one of the triangles of which the vertex is a member, each of these triangles is examined (Fig.6.10) to find this closest point.

The defined function f is a nonlinear function depending on $6 + t$ parameters.

In theory, one could simultaneously optimize the rigid and the nonrigid parameters. However, experiments have shown that it is more efficient to adjust them sequentially. Given an estimate for the pose parameters $\mathbf{R}, \mathbf{T}$ by applying the ICP algorithm (see section 4.4) the deformation parameters $w_1 ... w_t$ are adjusted using the Levenberg Marquard Method [PFTV92]; the partial derivatives of the objective function with respect to each of the deformation parameters can be computed analytically. The following overall strategy is applied to avoid local minima: Beginning with the first (most significant) deformation mode, the number of modes used to fit the data is increased successively in each iteration until the chosen maximal number of modes is reached. To decrease the search space, bounds to the parameters are applied. The change in rotation is limited to $\pm 30°$ around the initial guess and, the change in translation is limited to $\pm 10\%$ of the size of the object after the initial rigid registration. The allowed change in deformation is limited to $\pm 3\sqrt{\lambda_i}$,

where λ_i is the i^{th} eigenvalue of the corresponding mode.

6.5 Results

Ten right femurs have been used to build the statistical shape model as explained earlier in this chapter. Then both, simulated and real data have been used to carry out experiments.

Table 6.1 shows the relative importance of the modes of variation for the computed statistical knee model. It can be seen from Fig.6.11 that the first four modes already represent more than 90% of the shape variation in the model.

Nr.	Eigenvalue	Percentage
1	7375.37	75.60
2	814.50	8.35
3	530.53	5.44
4	218.55	2.24
5	199.29	2.04
6	165.28	1.69
7	153.29	1.57
8	132.40	1.36
9	101.81	1.04
10	65.54	0.67

Table 6.1: Relative importance of the modes of variation for the model

Excluding each of the other remaining 10 femurs from the 11 specimens, 10 further models can be computed. Table 6.2 shows the relative importance of the modes of variation for the knee model built while excluding the fourth specimen, which was recognized to be one of the smallest femurs after careful visual examination. As can be seen from the table, this results in a somewhat smaller relative importance of the first deformation mode describing the global scaling, but the general distribution of the eigenvalues does not change greatly. Computing the other 9 models gives similar results. The results presented in section 6.5.1 refer to the model built of the first 10 femurs.

Figure 6.12 shows the effect of applying ± 3 standard deviations of the first our modes of the obtained model to the mean shape. Mode 1, which accounts for almost 70% of the total variance within the model, can be seen to be primarily concerned with describing global scaling.

6.5.1 Simulated Data

Experiments with simulated data have been performed following the scheme shown in Fig. 6.14. To simulate the intra-operative data acquisition, the triangulated surface of the remaining femur (*test shape*) has been arbitrarily resampled. Afterwards several point sets each containing different numbers of points all lying in the accessible area on the femoral

Nr.	Eigenvalue	Percentage	Sum
1	5500.59	69.40	69.40
2	880.37	11.11	80.50
3	548.97	6.93	87.43
4	227.88	2.87	90.30
5	191.34	2.41	92.72
6	155.18	1.96	94.67
7	128.95	1.63	96.30
8	107.22	1.35	97.65
9	104.08	1.31	98.97
10	81.91	1.03	100.00

Table 6.2: Relative importance of the modes of variation for the model built while excluding the fourth specimen

joint surface during surgery have been acquired interactively. After aligning these simulated *test data* manually with the mean shape, the above detailed optimization strategy is applied to minimize the objective function.

As expected the final root mean square error (RMS) between the test data and the registered model decreases using more deformation modes ($\sim 0.7mm$ using one mode to $\sim 0.4mm$ using 5 modes) while this is not always the case for the RMS between the test shape and the registered model. 60 data points were used for this experiment, resulting in an execution time of approximately 8 seconds using 5 deformation modes.

Fig.6.15 (top) shows the model (triangle mesh) and the test femur after the initial rigid registration with the ICP algorithm. Fig.6.15 (bottom) shows the model and the test femur after the non rigid registration. The white spots represent the data points. One may observe that the fit between the test femur and the model also increases in those areas where no data points are available.

This is confirmed when computing the distance between the reconstructed model and the original test surface, for each point of the surface, thus providing spatial error distribution. Fig. 6.16 shows an error histogram and visualizes the error after the initial rigid registration (top) and after the nonrigid registration (bottom) by assigning a color to each surface point representing the local error.

Tests including different simulated data sets as well as different initial rigid body transformations lead to similar results.

6.5.2 Experiments with real intra-operative data

For two clinical cases, the surgeon acquired about 100 points randomly distributed on the femoral notch surface (not limiting the acquisition to a small patch and the anterior border of the notch). For these cases, it was possible to test the proposed statistical model fitting.

Results are provided in Table 6.3. Fig. 6.17 shows the deformed statistical model (tri-

angle mesh) that fits the points (black spots) collected during surgery (case 1). Although additional local deformation of the model would be necessary for perfect fitting, the result seems satisfactory. However, these tests do not evaluate the method's extrapolative power since the real shape is unknown, they only demonstrate two points:

- A sufficiently large collection of points is possible during surgery (compared with the necessary area defined by simulation tests)
- Interpolation by this technique provides sufficient accuracy for visual feedback

To evaluate the extrapolation quality for clinical cases, the deformed model would have to be compared with the complete shape of the femur from which the acquired 3D data originate. This could be done using a CT scan for instance.

RMS error	Case 1	Case 2
Rigid alignment	2.23	2.82
Model fitting with 2 modes	2.07	2.17
Model fitting with 4 modes	1.75	1.90
Model fitting with 6 modes	1.61	1.83

Table 6.3: Residual fitting errors (in mm) for two clinical cases using rigid alignment and model fitting with 2, 4 or 6 modes.

6.6 Discussion

It is considered that the results presented in this chapter show that the assumption of shape stability of the femur seems to be valid. Obviously, this needs further investigation, with particular attention to pathological deformations of the knee. For fresh ruptures of young athlete knees, it is reasonable to assume that the shape of the knee surfaces are normal, which corresponds to most of cases of ACL reconstruction. It is likely that pathological cases can be captured using a few additional statistical modes because the deformations often originate from the same problem and create the same effects.

The conducted experiments show that reconstruction from sparse data is possible using few modes of the model. This validates the statistical approach in the particular instance of the femur although it needs to be confirmed on a larger population. Using this method, it still needs to be investigated how many points and at which location are necessary for correct reconstruction. As a side effect of the method, the model can interpolate the data in a more robust manner than a single spline, without any topology consideration.

Actually, one major issue of the method is to know if a few data points (and at what location) are sufficient to capture the rigid body pose of the model, with an accuracy suitable for the application. This is highly dependent on the shape complexity and the number of modes necessary to capture the deformation. Clearly, a roughly flat surface with a high number of statistical modes cannot be reconstructed by this method. If too

many modes are used, there is a risk that extra-modes are used to compensate rigid body motion partially. Therefore, the method is suitable in two cases:

- The rigid body pose of the model is known independently of the method using least-squares fitting of anatomical or functional landmarks.
- The shape is reasonably complex and its deformations can be captured by using a low number of modes.

The second case is validated for the femur, on the basis of the experiments presented here.

Further investigations must show whether the accuracy of the model is sufficient to outperform the bicubic spline in interpolating the data points in terms of a more accurately computed surface. However, the method is not intended to cope with fine adjustments. Therefore, if accuracy is a severe requirement and strict interpolation of data points is mandatory, then it is necessary to perform additional unconstrained deformation of the model, for instance an FFD based volumetric deformation method that has been used in chapter 5.

6.6.1 Global Scaling

In the presented approach the effects of global scaling are captured by the statistical analysis rather then applying a scale normalization before performing the PCA as it is common. This enables the model to capture shape changes correlated with scale. On the other hand it biases the results to modes of the larger shapes. Further investigations must show which approach allows better shape recovery.

6.6.2 Tibia Model

The method would have probably difficulties in the case of the tibia considered alone. As less intra operative data are available for the tibia than for the femur, one can build one model including both tibia and femur taken in the blocked extension position of the knee. This seems a reasonable approach because the shape of femur and tibia belonging to one knee joint do not vary independently.

6.6.3 Improving the initial rigid registration between the intra operative data and the model

In the current implementation, a first rough alignment between the data points and the model is performed interactively by the user, followed by rigid registration using the ICP algorithm. For further automation, investigations must show how far other reference features can guide the matching process. Although the flexion-extension axis of the knee joint does not remain constant over the range of motion, it nevertheless could for instance serve as a functional reference feature similar to the concept of functional points described in 4.2.2.

6.6.4 Establishing the minimum number of points that meet the accuracy criteria for the application

Currently about 60 data points are acquired during surgery more or less arbitrarily distributed over the accessible area. As the operation time shall be kept as short as possible, it is desirable to acquire as few data points as possible while still meeting the accuracy criteria for the application. Obviously, not only the amount of data points but also their distribution over the surface, heavily influences the achievable accuracy. See, for instance [SHK95], for more about this (geometric constraint analysis).

6.6.5 Other Applications

The method can be applied to reconstruction of surfaces for which the shape is reasonably stable. For instance it can be applied to reconstruction of human faces using range images, or reconstruction of vertebrae using direct digitization of the posterior surface of the vertebra during surgery (see section 7.6), etc. Applications outside the medical world are also possible, for example in cases where it is necessary to digitize the surface of several instances of the same object. Not only the reconstruction is robust, but the model can also be used to infer some properties to the data. Labeling a region is thus automatic.

6.7 Conclusion

The main contribution in this chapter concerns a new method for inter and extra polating sparse scattered point data acquired on an anatomical surface. It is based on the fitting of a statistical shape model to the data points. First experimental results show that the chosen approach may be successfully applied to a system for computer-assisted anterior cruciate ligament reconstruction, although it must be mentioned that the 10 training shapes from which the model has been built are not sufficient to represent the natural shape variation of the femur. Therefore further experiments with a larger shape data base must be performed to further validate the chosen approach.

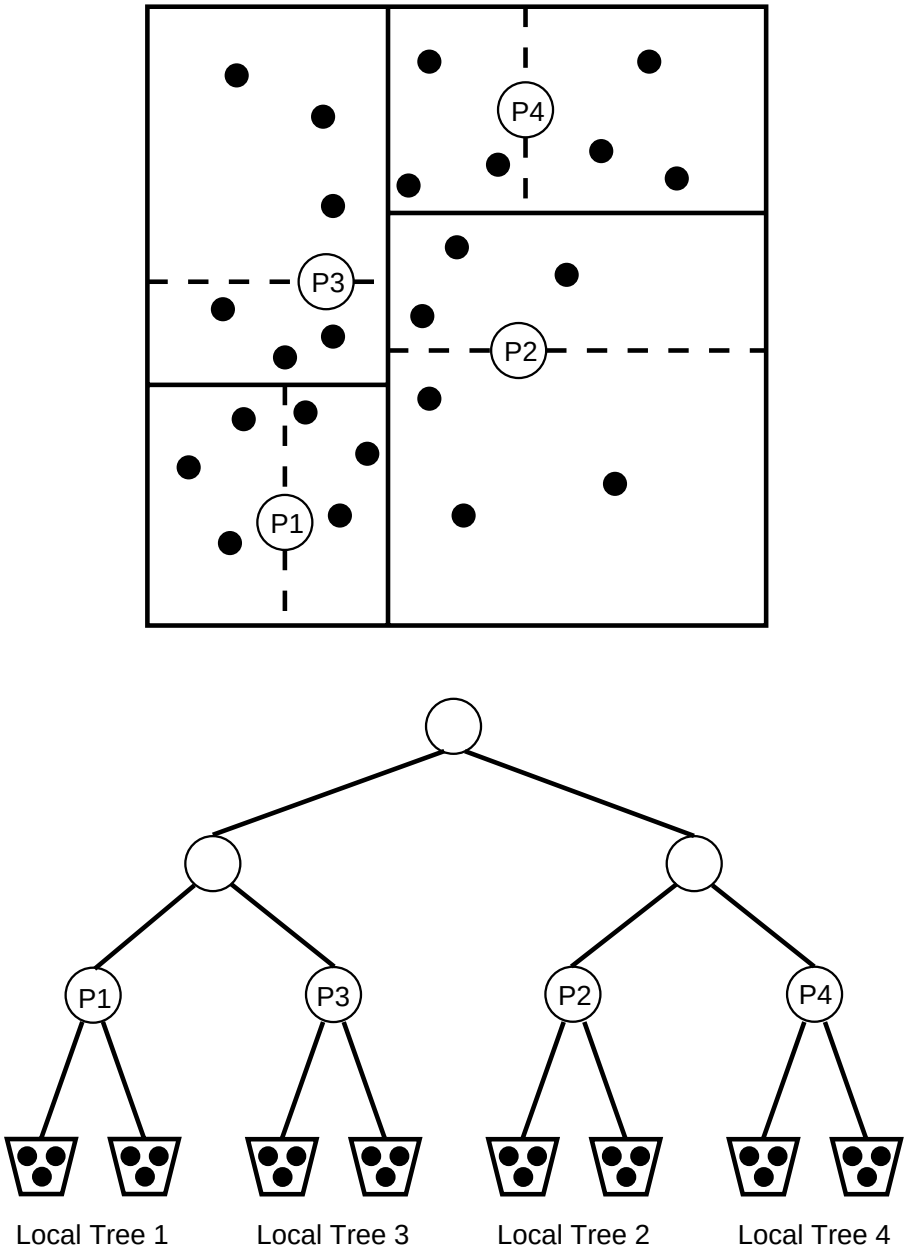


Figure 6.9: The two-dimensional k-d tree

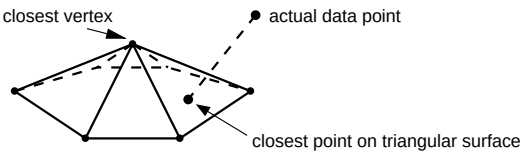


Figure 6.10: Searching the closest model point on the triangular mesh surface

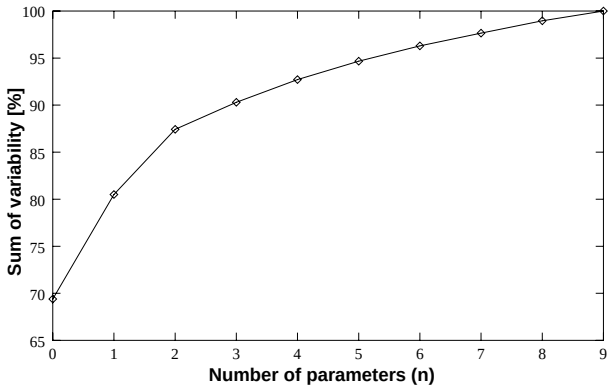


Figure 6.11: Captured variability of the statistical model of the distal femur as a function of the first n eigenmodes in percent

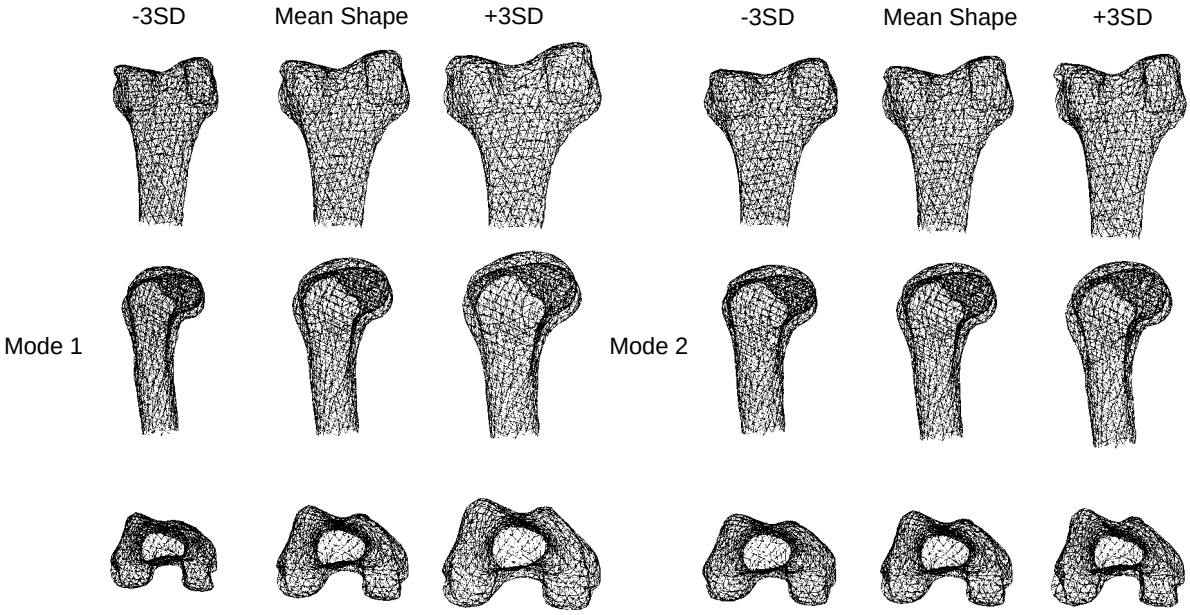


Figure 6.12: Applying 3 standard deviations of the first and second deformation mode on the mean shape

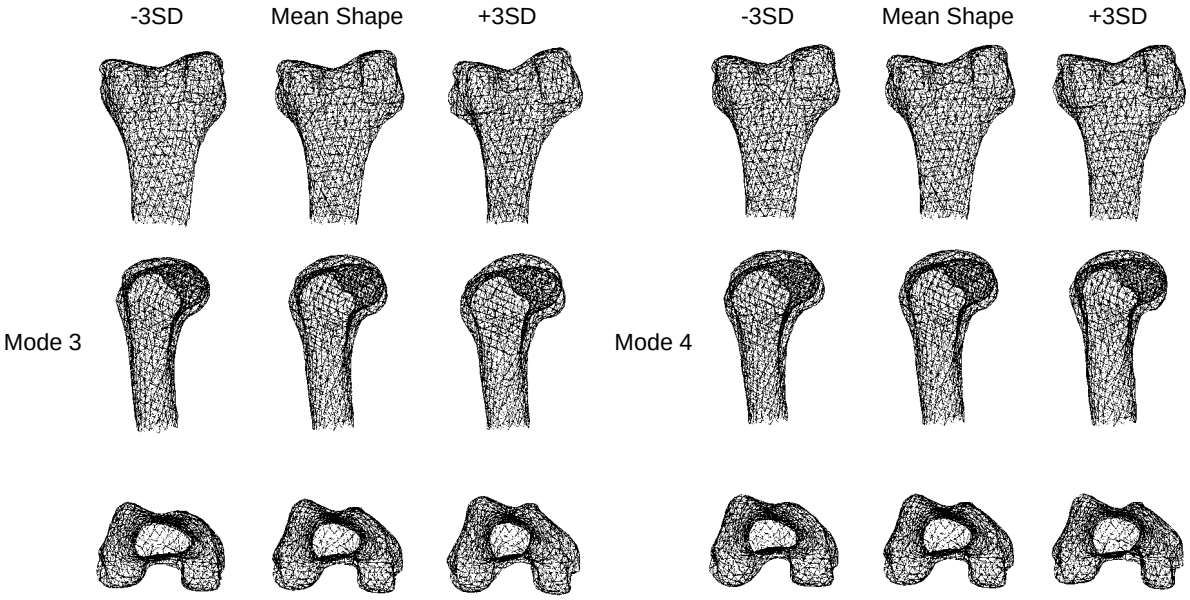


Figure 6.13: Applying 3 standard deviations of the third and fourth deformation mode on the mean shape

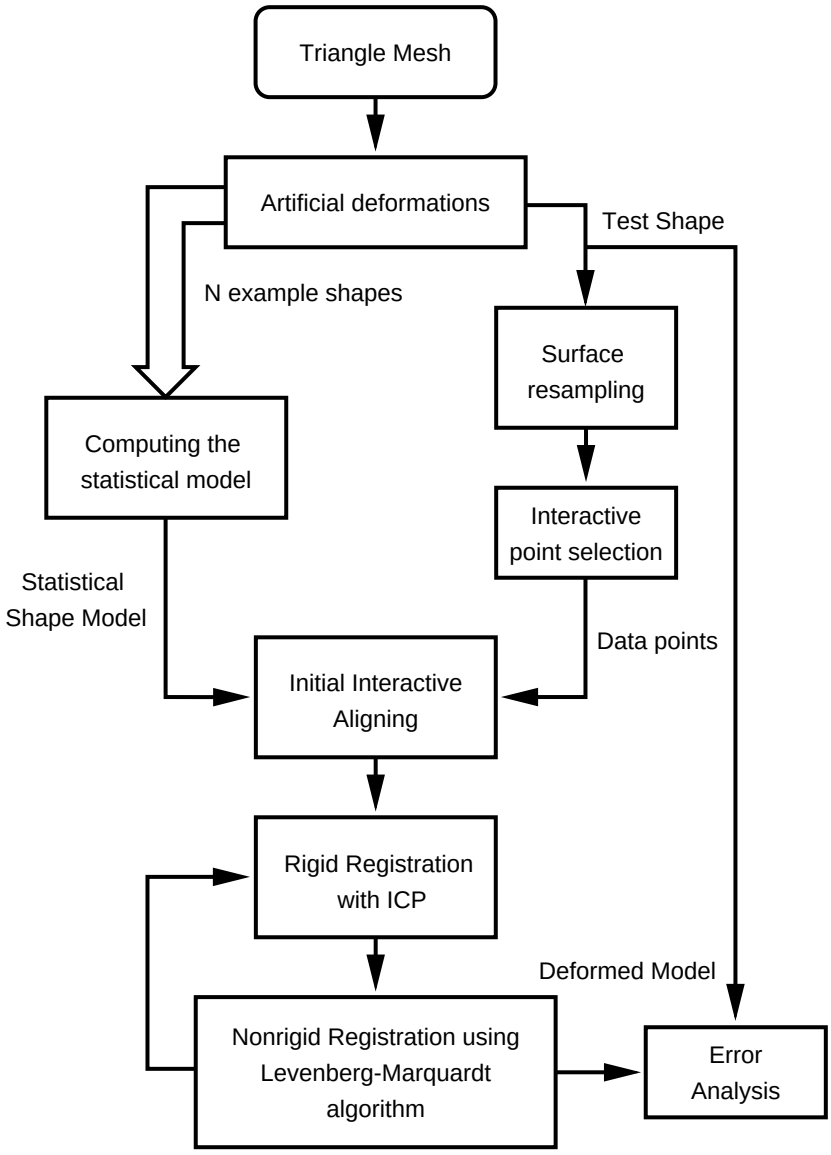


Figure 6.14: Flow chart of the performed experiments

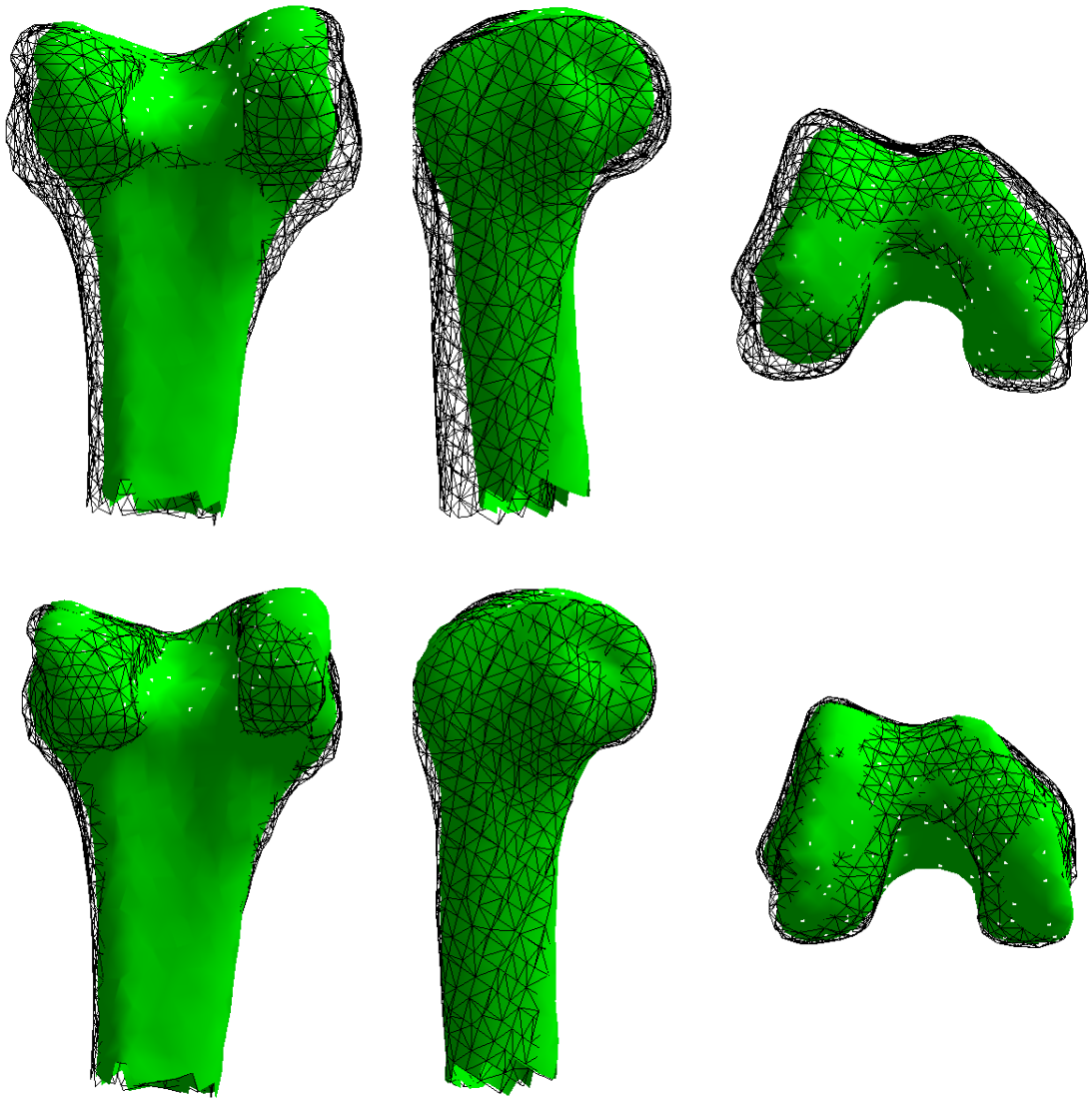


Figure 6.15: The model after rigid (top) and after non rigid registration (bottom)

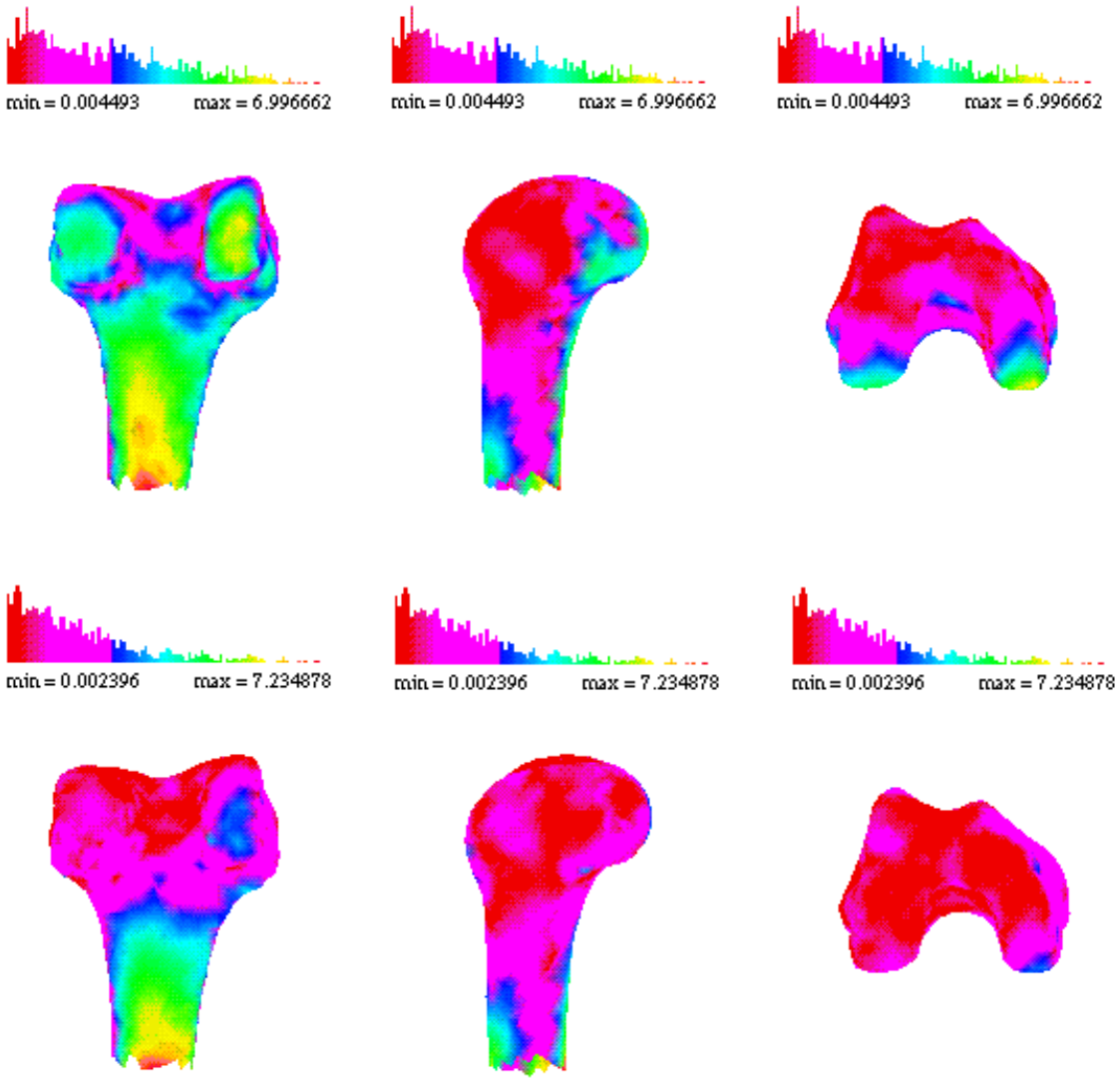


Figure 6.16: Error visualization after rigid (top) and after nonrigid (bottom) registration

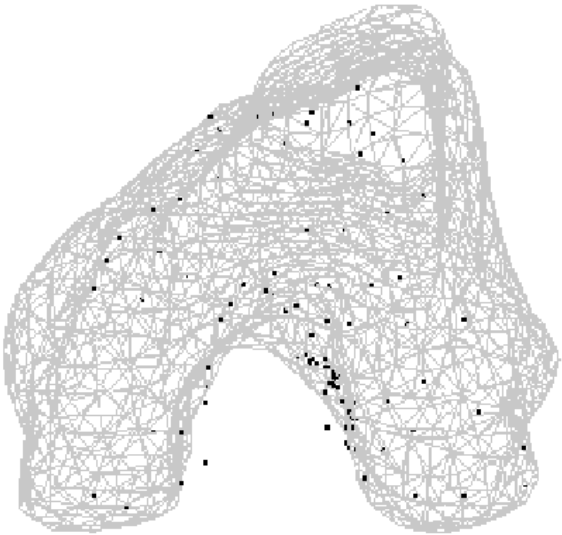


Figure 6.17: Clinical case 1: the model after non-rigid fitting using 6 modes in comparison with the sparse set of points.

Chapter 7

Nonrigid 3D/2D registration of a statistical shape model with few X-ray images

This chapter investigates how a statistical shape model can be used for nonrigid 3D/2D registration with very few X-ray projections. X-ray images are the dominating image modality in the operating room. Despite its widespread acceptance and utility, however, fluoroscopy is associated to several disadvantages. One important issue is the radiation exposure, particularly to the surgeon's hands. Recent data suggest that spinal surgeons, in particular, are at significant risk for fluoroscopy-related radiation exposure [ea, ea93]. The most obvious disadvantage perhaps is the fact that only 2D images are available. More precisely only one single real-time planar view is usable at any given time. Consequently, for procedures requiring multiplanar fluoroscopic visualization, the C-arm has to be repositioned throughout the procedure. This process is often tedious, time-consuming, and frustrating. Under ideal circumstances the surgeon holds an instrument perfectly still in one plane while correcting its position in the other. The X-ray technologist efficiently repositions the C-arm to obtain the perfect view in each desired plane while maintaining ideal sterility and without ergonomically challenging the surgeon with the C-arm.

Thanks to his anatomical knowledge the surgeon is used to mentally fusing 2D images taken from different view points. However for many applications this mental registration is not sufficient to obtain all necessary information about the anatomical situation in order to properly perform the surgery.

In many applications of surgery such as orthopedics for instance, it is desirable to define a surgical planning on 3D images and then to execute this plan on the patient. Therefore, since the advent of Computed Tomography, various surgical interventions are preceded by the construction of a CT-based 3D model of the object of interest to provide the surgeon with spatial information which is lacking when using only 2D images. Since the introduction of computer assisted surgery systems it became possible to register these pre-operative acquired CT data with intra-operatively acquired X-ray images using fiducial based or surface based registration methods (see section 4.2), thus unburdening the surgeon from fusing the images with the patient mentally.

But CT based CAS systems are associated with several drawbacks: First, the total X-ray dose for the patient raises significantly by the CT exam. Further it significantly increases the overall intervention costs as well as its duration due to the necessary registration procedure. Furthermore, for many applications the computation of a detailed (though very precise) 3D attenuation map is not mandatory, i.e. reconstruction of the organs shape would be sufficient. Therefore it is desirable to be able to infer 3D-information from the intra-operatively acquired 2D X-ray images to facilitate the navigation within the patient and thus allowing to abandon CT data acquisition at least for many standard surgical applications, where this data is not otherwise required for diagnosis purposes. Abandoning the CT scan also eliminates the often time consuming registration process between pre-operative and intra-operative images.

By combining current C-arm fluoroscopy with computer-aided surgical technology, many advantages of fluoroscopy can be enhanced, while minimizing or eliminating its disadvantages. In [Hof97] for instance, authors propose to acquire several images using a classical C-arm equipped with an image intensifier and to track the position and orientation of the surgical tools, the image intensifier and a patient's reference with an optical localizer, thus allowing to compute relative movements of the patient or the surgical tools with respect to each acquired image during the intervention. Although this system is a considerable improvement, real 3D information is still missing. Recent publications ([ea01b]) at least raise doubt about sufficient obtainable precision for these so called "virtual fluoroscopy" systems for pedicle screw placement in the case of scoliosis for instance.

The objective of this chapter is to explore a method to recover the 3D shape of the patient bones or organs intra-operatively using a very limited number (2 - 3) of calibrated X-ray images. This is accomplished by deforming a statistical 3D model to the contours segmented on the X-ray images. Here work is concentrated on the application of this method to bone reconstruction. The algorithm starts from segmented contours of the bone on the x-ray images and an initial estimate of the pose of the 3D model in the common coordinate system of the set of X-ray projections.

Fitting the model to the contours is then achieved by using a generalization of the iterative closest point algorithm to nonrigid 3D/2D registration, following the concept of section 4.4. As a matter of course, only healthy organs or shape pathologies which are possible to capture by statistical analysis of a population, may be reconstructed with this method; e.g. fractured organs can not be handled.

The remainder of this chapter is organized as follows: Section 7.1 provides a brief overview. Section 7.2.1 presents a generalization of the Iterative Closest Point algorithm for contour based 3D/2D registration. The next section shows how to efficiently compute matched point pairs by computing the model's contour generators. Subsequently it is shown how to fit the model to the projection data. Section 7.4 provides results obtained with simulated and real data.The experiments carried out are based on segmented X-ray images; in this dissertation work is not focused on the vast field of 2D segmentation methods. However, section 7.5 briefly presents some preliminary results obtained with a model based approach combining a low level pre-segmentation of the X-ray images with model based approach, as proposed in [Ion98]. Section 7.6 presents the results of an

experiment combining the methods proposed in this and the previous chapter, constituting a hybrid registration approach relying on both 2D projective and 3D point data. Section 7.7 summarizes this chapter.

Image intensifiers, are subject to geometric distortions due to non planar shape of the image intensifier and external magnetic fields (see section 3.1). Calibration techniques such as the NPBS method [CLSC92] for instance can be used to correct these distortions as well as to compute a pseudo focal point of the source. This calibration is not further addressed here. However full digital X-ray detectors without any geometric distortion [CCD+98] begin to appear on the market and are likely to replace the image intensifiers in the future.

7.1 Related Work

7.1.1 X-Ray Tomography

In (X-ray) absorption tomography, the selective absorption of X-ray photons by the different tissues being imaged is used for the reconstruction of a density image from a complete set of projections (complete with respect to the Shannon sampling theorem) on the grounds of the Radon theorem [Nat86]. Reconstruction algorithms are usually either based on filtered back projection methods or algebraic reconstruction techniques (ART) [KS87].

Common CT-scanners are based on 1D X-ray detectors thus providing a stack of reconstructed 2D images (fan beam projections). However 3D tomography systems (based on cone-beam projections) begin to appear on the market [ea98]. There are also attempts to use intra-operative fluoroscopy systems for Computed Tomography [Eea00]. Although promising results are reported, this is associated to several problems. One serious problem is the mechanical instability of the C-arm. Although it is possible to construct more stable C-arms (up to a certain extent, since the C-arm must still fit the OR and must remain ergonomic) most known reconstruction algorithms require a specific trajectory.

One objective is to reduce the dose delivered to the staff and the patient. Contrary to the patient, the staff can be protected from the radiation during intra-operative image acquisition. For good image quality, tomography usually requires several hundred projections. Fig. 7.1 shows reconstructed images of the lumbar spine when vigorously reducing the number of used projections. With less than 16 projections image structures become impossible to recognize. It is reminded that the objective is to obtain acceptable shape reconstruction using 2-3 projections only.

Boundary detection

As said before, for many medical applications such as orthopedics, determination of the object shape i.e. boundaries (e.g. bone surfaces) is sufficient to perform the task. In many cases it is not necessary to compute the 3D attenuation map.

Therefore several authors addresses the problem of segmentation of objects based on the sinogram, i.e. without prior reconstruction. In [Thi91] authors propose to reconstruct

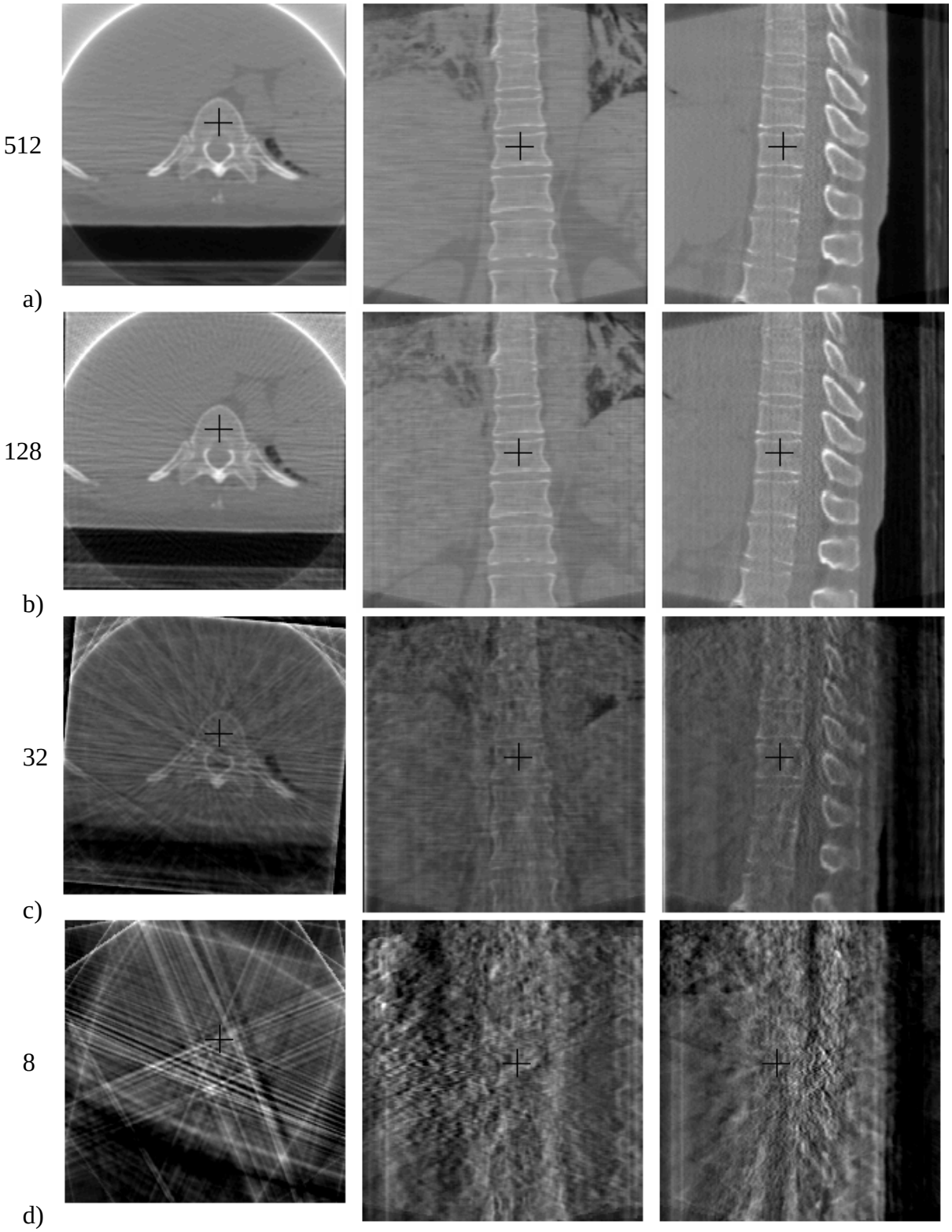


Figure 7.1: Spine images reconstructed with Feldkamp algorithm using different number of projections. The projections were forward-projected from a regular CT scan (Visible Human Project, VHP). The original image size is 512^2 with $0.9375mm^2$ pixel size, slice thickness 1.0mm. The projection image size was 300^2 with $1.43mm^2$ pixel size. (Courtesy of Xuan Liu, Brussel University, Belgium) Number of projections used for reconstruction: a) 512, b) 128, c) 32, d) 8

object boundaries from raw CT data without first computing the attenuation map, but their algorithm requires a dense set of x-ray projections. In [KCD98] the former approach is applied to non destructive testing (NDT), extended by incorporating a priori knowledge in the form of parameterized object models, which are detected using the Hough Transformation. But the models are limited to simple objects like ellipses. In [Ela01, ER01] the use of a level-set surface model in the context of segmentation of spiny dendrites from limited angle electron microscope tomographic data, segmentation of vascular structures from simulates sparse-angle angiography data and segmentation of lungs and the torso from simulated transmission scans is investigated. The level-set approach is a rather low level model; it provides topological flexibility but may not be contrained enough to cope with very few projections and noisy data. In [BBLRT99] authors use a free form deformation model to reconstruct shapes of constant interior density in the context of SPECT imaging. Robustness to noisy data is shown, but the effect of incomplete data is not investigated.

Ordinary x-ray tomography is global, i.e., reconstruction at a certain point requires integrals over lines far from this point. In cases, where the region of interest (ROI) is much smaller than the whole object, local tomogography methods can be considered to find discontinuities (object contours)in the ROI [FKH+01]. Local tomography is based on X-rays passing through the ROI only and thus helps reducing patient X-ray exposure. Due to the non-local nature of the Radon transform, a perfect reconstruction of the attenuation map is impossible with only local projections data.

When the number of available projections is small, image reconstruction becomes an extremely ill-posed problem that can be (approximately) solved only with the formulation of a set of strong assumptions about the structures in the scene.

Discrete Tomography

One possibility is to consider regularization by limiting the solution space to a reduced set. For instance, constraining the solution to be binary is a possibility (discrete tomography). Examples arise e.g. in NDT, where a common problem is to detect and identify fabrication defects in homogeneous industrial pieces. In [VD97] authors take into account a small number of cone-beam projections (3-7 projections) considering a priori knowledge about the object given in general form. Results are presented for binary objects, but the authors claim that the algorithm is extensible to objects with multi-level structure. In [Kli90] authors present an iterative two step algorithm for 3D reconstruction using a limited number of projections (3). The first step of the algorithm is a classical iterative reconstruction ART type method which provides a rough volumetric reconstructed 3D zone containing a flaw. Then in the second step this reconstructed zone is modeled by a Markov Random Field (MRF).

In the medical domain binary reconstruction arises for instance when reconstructing a vascular tree from a small number of radiographs using a conventional angiography device. The acquisition is performed before and after injecting a contrast agent in the arteries, which may be assumed to provide a homogeneous medium. When the two projections are subtracted, the problem is equivalent to reconstructing a binary image where zero corresponds to the background and one to the contrast agent, thus determining the geometry of the vascular tree. In [LBM+96] authors present a method for cerebral vascu-

lar reconstruction from six projections. The approach is based on a regularization of the inverse problem by a smoothness function using the discrete smooth interpolation (DSI) formulation. In [RTAR+95] authors present a method for reconstruction of binary objects from two orthogonal projections using graph theory and its application to dentistry. In general these methods aim more at reconstructing the topology of objects than their more or less precise shape. Considering vessels for instance it is sometimes more important to know about all the branchings than the exact shape of the vessel. Unfortunately the conditions allowing to make the assumption of binary circumstances are not given in the case of orthopedic, because there is not enough contrast between the object of interest and the surrounding tissues.

7.1.2 Model based 3D Surface Reconstruction Techniques Using few 2D X-Ray Projections

Before discussing the more general shape fitting problem different concepts for rigid registration are reviewed.

Rigid 3D/2D Registration (Pose Estimation)

Basically two different concepts can be distinguished:

Contour based Registration One type of algorithm is based on contours and requires prior segmentation of the object in the 3D-image as well as in the 2D-image [LS95, FAB94, Gue98] although in [HSLC95] authors propose a cooperative approach between registration and 2D segmentation.

Intensity based Registration The other concept does not need segmentation and is based on comparing the gray value distribution of the 2D-image with the projection of the 3D-image under current registration parameters into the image plane (Digitally Reconstructed Radiographs, DRR) [LFK94, GDT+96, RBB+99, GBB96]. Due to the high computational cost for computing the DRR, this method is rather slow although in [Wee99] authors recently proposed a promising technique for considerable acceleration by using the shear-warp factorization [LL94]. Authors claim to achieve full 6 degree of freedom registration of a CT model of a vertebrae within 4 seconds. Currently low cost hardware architectures are under development which achieve fast ray casting (up to 30 frames per second) of rectilinear grids using this acceleration technique [RPSC99]. But so far only parallel projection is available. Other techniques use rather standard graphics hardware (2D or 3D texture mapping hardware) to speed up the volume rendering process [WE98, DKC+98]. See [PWL+98] for a comparison of similarity measures used for intensity based registration.

A typical assumption formulated when a very small number (2-3) of projections is available for shape reconstruction, is that only a single structure is present in the scene. The additional hypotheses that the structure has a constant density, that the rest of the imaged volume is empty and/or that the cross sections of the structure respect predefined

constraints allow the reconstruction of such a structure. The strong assumptions described above prevent the applicability of these methods to most anatomical structures of diagnostic interest. In general neither the single object nor the constant density assumption is valid in the case of bone reconstruction for example. Abandoning these restrictions needs more a priori knowledge.

3D Surface Reconstruction based on projections and Deformable Surface Models

One of the first works concerning the reconstruction of a 3D object from silhouettes was proposed by Martin and Aggarwal [MA83]. They produce polygonal 3D objects backprojecting the silhouettes. Benjamin et al. [Ben95] generate surface information using common points, present in different projections, and the tracks between the points, extracted from about 10 projection images. Caponetti and Fanelli [CF93] present a method to extract 3D geometry of bones from two orthogonal projections. The initial estimation of the 3D bone structure, produced by backprojected profile points, is refined by B-spline interpolation. In [Nik96] authors reconstruct femurs from 2 orthogonal X-ray images. They separate the femur into 3 subparts each of them assumed to be round. They fit cubic parametric surface patches to the subparts and then assemble them to a complete model. Kita [Kit96] developed a method to analyze X-ray images of the stomach using a deformable 3D model. The model is a tube which is first initialized using only one projection. Afterwards, the model is deformed using the other projections. Theoretical work concerning surface reconstruction from 2D silhouettes has been discussed in [Lau95]. Lötjönen et alt. [LMNK99] use a geometric prior model of the lung to reconstruct the organ from manually segmented 2D X-ray projections. In [HL96] authors present a method for constructing a 3D individualized head model by fitting a generic head model to facial features extracted from the side and front views of the head. Based on local maximum curvature tracking, corresponding point pairs are found on the side and front view and their 3D location is used to compute a distortion vector field for specializing the generic model for the input images. The work of [PV97] is one example where authors aim to recover shape from one single X-ray image by exploiting both, geometric and densitometric constraints while making two assumptions: the density of the structure to be recovered is approximately constant and the surface of each structure is smooth. This approach shares ideas from the work of [TWK88] where authors recover the 3D shape from the 2D profiles of an object using a deformable tube and a deformation technique which is controlled by physically based intrinsic and extrinsic forces. Both geometric and intensity information are also used in [SNP95] where authors reconstruct 2D objects from 1D projections using deformable spline curves.

Deformable models require an initial position in order to converge properly to the desired solution. This is generally achieved by a suitable initialization (see section 4.2.4).

7.2 Matching a statistical shape model with segmented X-ray projections

Here the shape recovery problem is formulated as a nonrigid registration between a statistical deformable shape model and the contour data extracted from the X-ray images.

7.2.1 The ICP algorithm for 2D/3D registration

Following the concept of section 4.4 it is possible to use a generalization of the ICP to perform the 3D/2D registration. Therefore it is necessary to define correspondences between the model and each projection ray $\mathbf{p}_{i,i=1...P}$ of the X-ray images, which are defined by the coordinates of the contour points x_i, y_i in the image plane and the focal point $\mathbf{f}$ of the source. This can be accomplished by associating the endpoints of that line segment originating on the projection ray and ending on the model surface such that their distance to each other is minimal. It is interesting to note that at each iteration step new points on both the model and the projection rays may be selected, while in the 3D/3D case (chapter 6) the points in the data set remain the same throughout all iterations. In [WH96] authors describe this approach for rigid 2D/3D registration of CAD models to video camera images but do not address the problem of quickly finding those points on the projection ray and the model that have the smallest distance to each other. This is a key step for applicability within intra-operative applications and is addressed in the next section.

7.2.2 Efficient matched point pair building

To efficiently find the above defined correspondence an approach described in [Gue98] is used. First the actual contour generators $\mathbf{g}_{i,i=1...G}$ of the model are computed. Contour generators are those object features constituting the (inner and outer) contours of the object in image space with respect to the current projection parameters. As the geometrical presentation of the model is triangle mesh, the contour generators are a subset of all triangle edges. Thus establishing correspondence results in the simple computation of shortest lines between two 3D line segments. When the model is perfectly aligned with the projection rays, the latter intersect those triangle edges previously found to be contour generators. In [Gue98] authors call the contour generators "apparent contours" and use them for a rigid registration algorithm to match a CT model with fluoroscopic images. In the chosen implementation, the triangles in the mesh are defined by pointers to an edge list. Each edge in the edge list points to the two vertices in a vertex list defining the edge. This representation allows to efficiently compute the contour generators using the following criterion: For each triangle the viewing direction is defined as the vector originating from the center of projection to the triangle centroid. If the triangle normal, defined by the cross product of ordered oriented triangle edges constitutes an obtuse angle with the viewing direction, the triangle is said to be visible and invisible otherwise. An edge is a contour generator if the triangle on one side of the edge is visible and the triangle on the other side of the edge is invisible. All edges meeting this criterion are stored in a list and

rather than performing brute force search within the complete edge list of the model the search only has to take place within this subset to find matched point pairs. For a model consisting of about 5000 edges for instance, there are only about 300 contour generators for each perspective projection. Fig 7.2 shows the correspondence between one projection ray and the current contour generators of a model of the distal part of the femur.

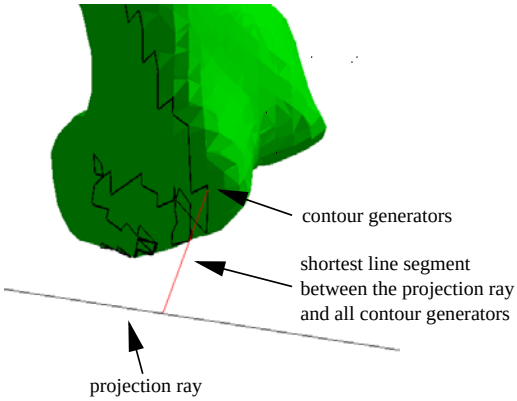


Figure 7.2: Correspondence between one projection ray and the current contour generators of a model of the distal part of a femur

7.3 Model Fitting

To recover the shape from the projection rays it is necessary to find the rigid transformation (rotation $\mathbf{R}$, translation $\mathbf{T}$) between the matched point pairs and the decomposition of the t preserved eigenvectors of the statistical model in such a way that the distances between them are minimized. The objective function to be minimized is defined as follows:

$$E(\mathbf{R}, \mathbf{T}, w_1 ... w_t) = \sum_{j=1}^{P} \min_{1 \leq k \leq G} \|\mathbf{p}_j - (\mathbf{R}\mathbf{g}_k(w_1 ... w_t) + \mathbf{T})\|^2 \tag{7.1}$$

As for the 3D/3D case in section 6.4 the rigid and the nonrigid parameters are adjusted sequentially. Given an estimate for the pose parameters $\mathbf{R}, \mathbf{T}$ by applying the generalized ICP algorithm, the deformation parameters $w_1 ... w_t$ are adjusted using the Levenberg-Marquardt algorithm; the partial derivatives of the objective function with respect to each of the deformation parameters can be computed analytically as in the 3D/3D case that was discussed in section 6.4.

The following strategy is applied to avoid local minima: Beginning with the first (most significant) deformation mode, the number of modes used to fit the data is increased successively in each iteration until the chosen maximal number of modes is reached. Bounds to the deformation parameters are applied to force the model to deform only in an anatomical reasonable range as in the 3D case.

7.4 Results

7.4.1 Simulated Data

Experiments with simulated data were carried out using the shape model of the distal part of the femur (section 6.5) and a statistical model of lumbar vertebrae. By means of a simulator tool rotation and translation parameters of a 3D model can be changed interactively and the contour generators can be projected on a virtual detector plane thus providing contours for known projection parameters. The experiments were performed using an image plane / focal point distance of 1000mm thus roughly approximating real conditions when using a C-arm.

Experiments with the femur model

Figure 7.3 shows 2 simulated X-ray shots taken for two approximately orthogonal view points around a model of the distal part of the femur (one lateral and one AP view).

Fig 7.4 shows the shape model before registration (a), after rigid (b) and after nonrigid (c) registration. One recognizes that the projection rays are tangential to the object surface after the nonrigid registration.

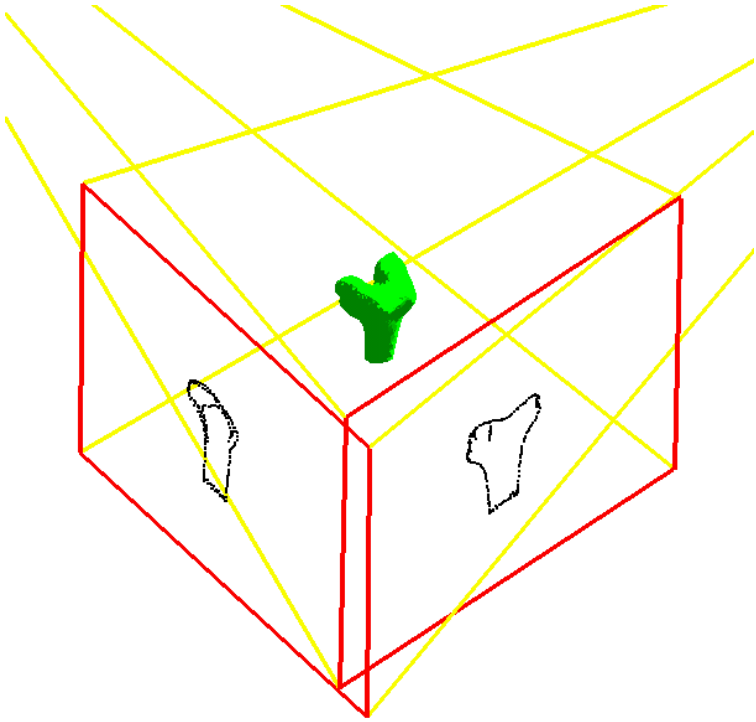


Figure 7.3: Two simulated X-ray shots acquired around a femur surface model

Experiments using different numbers of X-ray images show that within the current implementation two orthogonal views establish the best compromise between accuracy and computation time. Table 7.1 shows registration results for different numbers of calculated

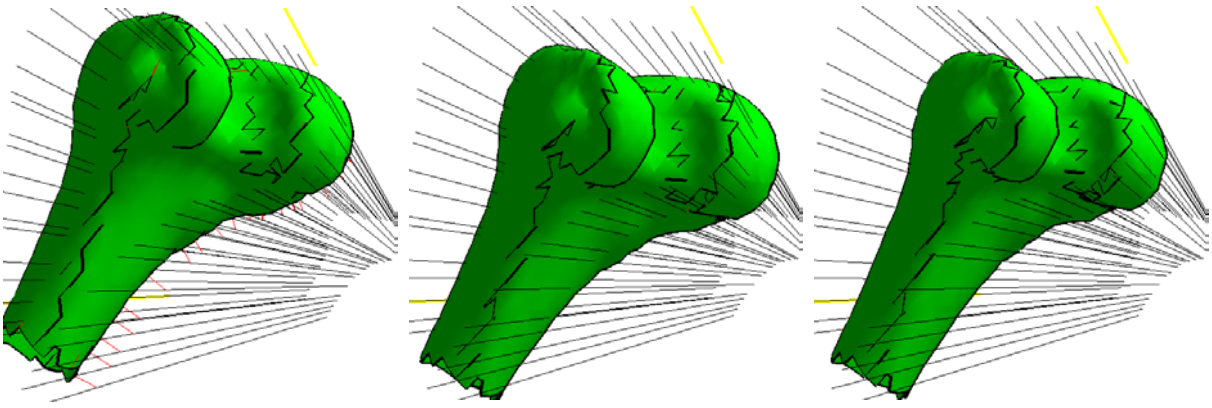


Figure 7.4: Surface model of the distal part of the femur: (a) before registration, (b) after initial rigid registration, (c) after nonrigid registration

projection rays per X-ray image. In this experiment two orthogonal X-ray views and four deformation modes were used. Shown are the RMS between the projection rays and the model (left column) and the RMS between the surface model of the shape to be recovered (reference) and the deformed model (right column).

	RMS (mm)	
rays	model-projections	model-reference
10	0.34	1.3
20	0.52	1.2
50	0.49	1.13
100	0.55	1.05
200	0.77	0.99

Table 7.1: RMS error for different number of projection rays per view (in mm)

Comparing the accuracy of 2D/3D matching algorithm with the 3D/3D registration algorithm presented in chapter 6 yields the following results: Approximately 500 points (randomly distributed on the surface of a CT-model to be recovered) were first registered rigidly with the mean shape, resulting in a RMS of 2.44mm. The nonrigid registration between the 3D data set of the test femur and the deformable model using 4 deformation modes results in a final RMS of 0.85mm. Using 2 (orthogonal) views, each with about 200 projection rays results in a final RMS between the deformed model and the shape to be recovered of 0.99mm, which is only slightly more with respect to the results obtained with the 3D data set.

Experiments with a vertebra model

A shape model of lumbar vertebrae incorporating 30 data sets was computed. Fig. 7.5 shows the effect of applying the first 5 deformation modes to the mean shape. Fig. 7.6 shows the captured variability of the statistical vertebra model as a function of the first n eigenmodes.

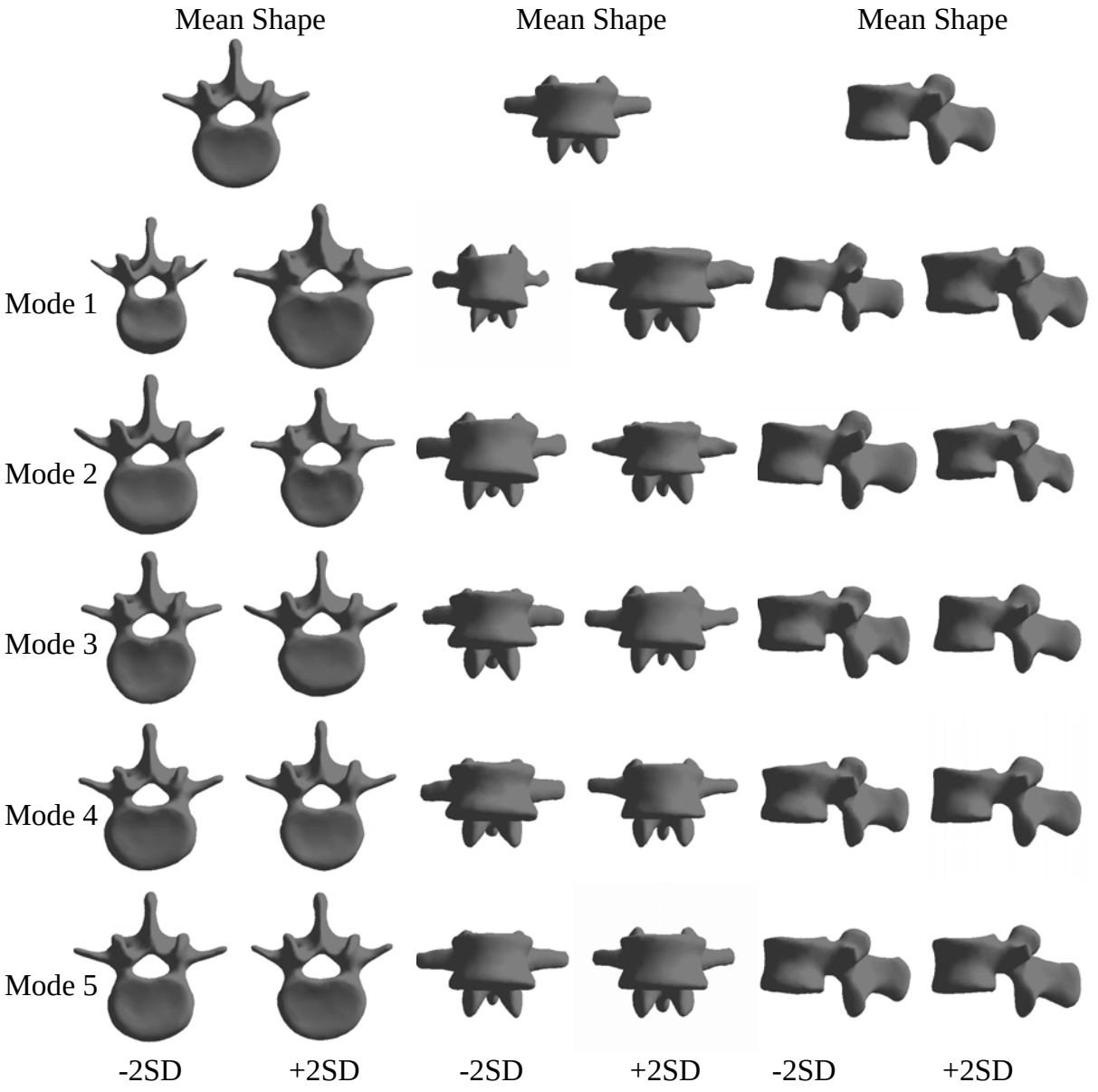


Figure 7.5: The first 5 deformation modes (covering 85 percent of the deformation contained in the 30 training shapes)

Experiments were performed using 2 orthogonal simulated X-ray views (Fig. 7.7) based on 10 different CT-models that were not contained in the population used to build the model.

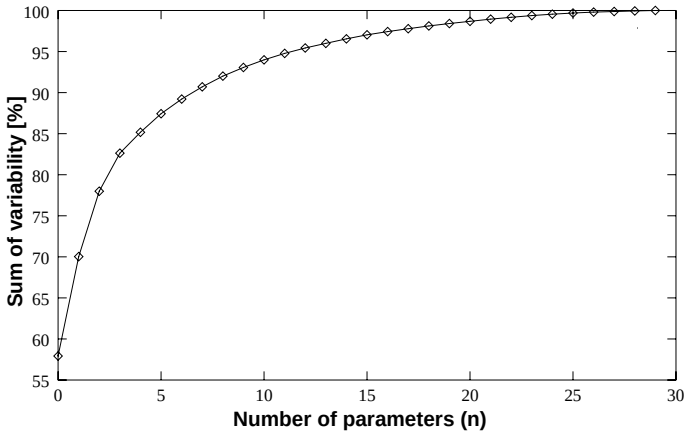


Figure 7.6: Captured variability of the statistical vertebra model as a function of the first n eigenmodes in percent

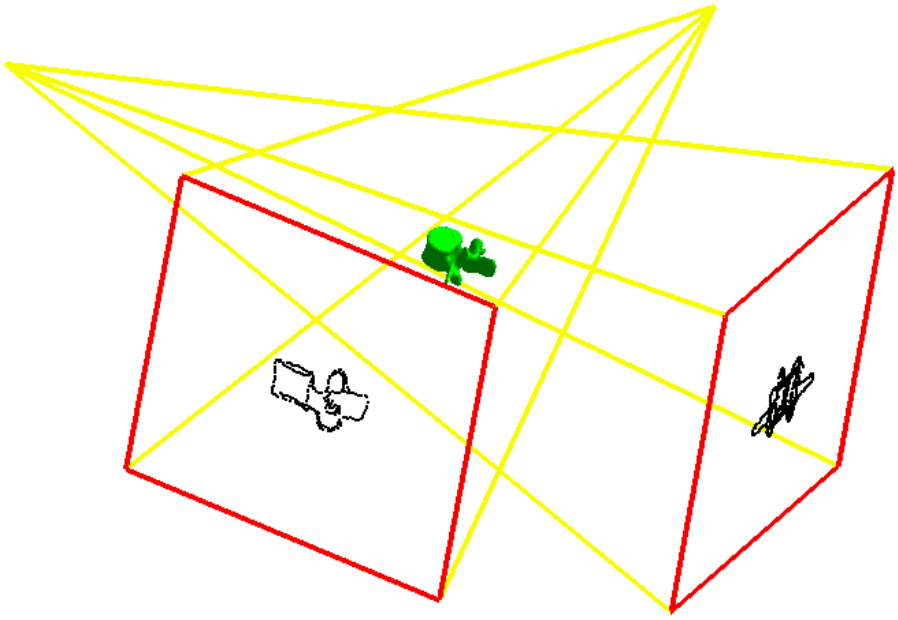


Figure 7.7: Reconstruction of a lumbar vertebra using 2 orthogonal x-ray projections

In each simulated projection 400 randomly chosen (internal or external) contour points were maintained thus taking into account the fact that in practice not all contour points will be reliably detectable (Figure 7.8 (b)). One additional data set was generated by deleting all internal contours taking into account that those are less likely to be reliably detectable in practice (Figure 7.8 (c)).

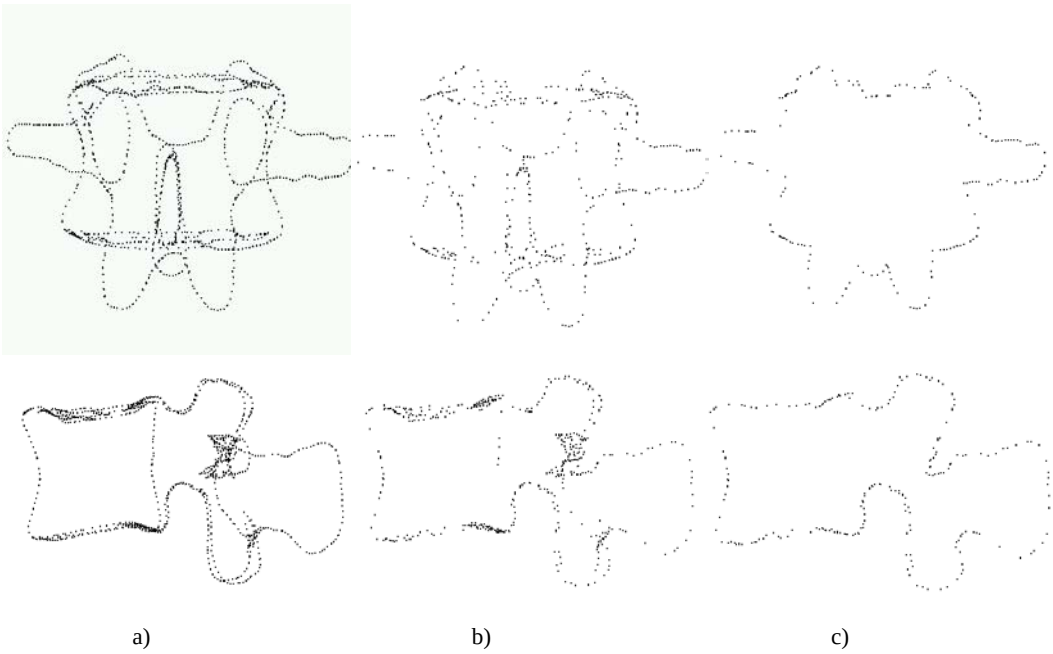
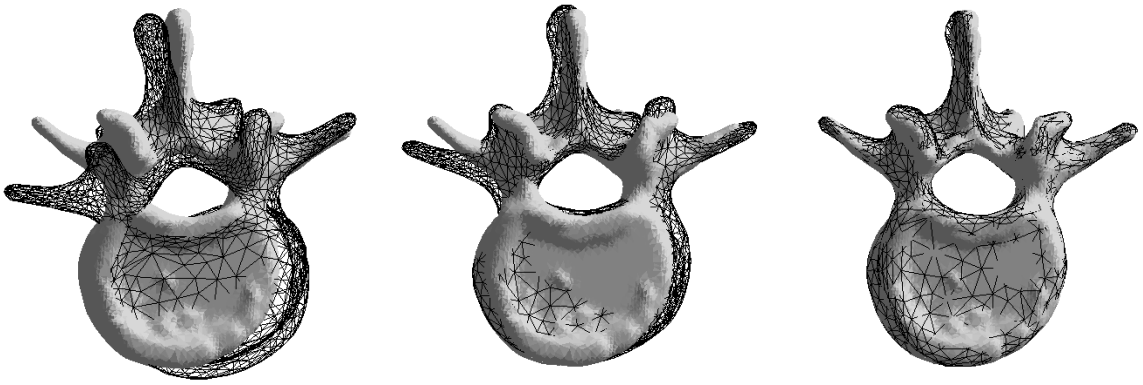


Figure 7.8: Contours for reconstruction using simulated projections: a) high density inner and outer contours, b) low density inner and outer contours, c) low density outer contours (top AP view, bottom lateral view)

Each of both low density datasets containing the two projections was then registered with the vertebrae shape model using all deformation modes following the strategy detailed in section 7.3. Fig 7.9 shows the shape model (mesh) after manual alignment (a), after rigid (b) and after non rigid (c) registration together with the CT-model from which the projections used for this experiment were generated. Fig. 7.10 shows the evolution of the RMS error between the projection rays and the model surface throughout the optimization process using increasing numbers of modes.

Table 7.2 summarizes the resulting RMS errors between the model and the projection rays and between the model and the underlying CT-model for the 10 cases using the data set containing only the external contours. The final average RMS between the deformed shape model and the underlying CT-model (reference) was 0.62 mm. The same experiment carried out using internal and external contours results in slightly lower final RMS errors ($\Delta \approx 0.1mm$) Fig. 7.11 visualizes the final RMS error distribution on the surface for one of the cases.One may observe that after non rigid registration only small surfaces patches of the model remain relatively distant to the underlying CT model surface; this is confirmed when comparing the error histograms after rigid and non rigid matching. Table 7.3 shows



a) b) c)

Figure 7.9: Shape model of a lumbar vertebra matched to 2 orthogonal X-ray views: a) after manual alignment, b) after rigid registration, c) after non-rigid registration.

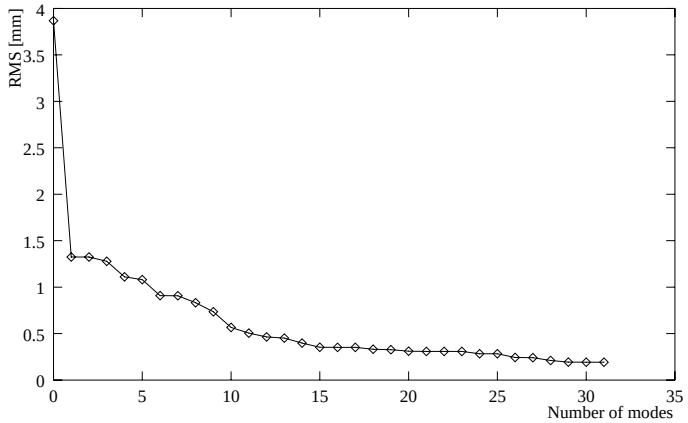


Figure 7.10: Evolution of the RMS error between the projection rays and the model surface using increasing number of modes for minimization. The first data point represents the RMS after manual alignment the second one after rigid registration.

the overlap measure between the shape model and the underlying CT-model after rigid and after nonrigid registration for each of the 10 cases; The average overlap between the deformed shape model and the underlying CT-model was 91.12 percent. Results are consistent with the RMS based results.

Nr.	rigid RMS model-projections	non rigid RMS model-projections	rigid RMS model-reference	non rigid RMS model-reference
1	1.29	0.35	1.69	0.44
2	1.44	0.19	1.89	0.36
3	1.52	0.22	1.81	0.29
4	1.32	0.31	1.78	0.42
5	1.49	0.50	1.88	1.32
6	1.49	0.18	1.62	0.38
7	1.32	0.36	1.91	0.63
8	1.20	0.45	1.78	1.29
9	1.37	0.29	1.77	0.54
10	1.44	0.32	1.73	0.56
mean	**1.39**	**0.32**	**1.79**	**0.62**

Table 7.2: Matching errors (in mm) for 2D/3D registration after rigid and non-rigid matching using two orthogonal projections containing the external contours of a vertebra

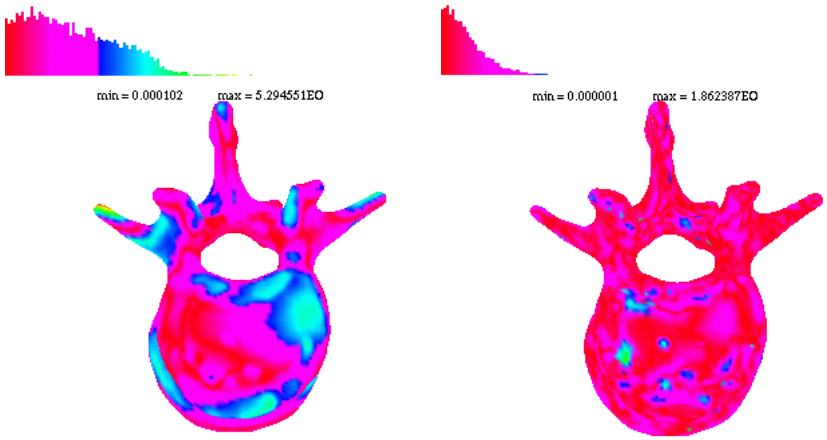


Figure 7.11: Error visualization for reconstruction of a vertebra using 2 orthogonal views

7.4.2 Real Data

To obtain more realistic results an experiment using real data was performed. A cadaveric lumbar spine was attached to a computer controlled turn-table and imaged twice for

Nr.	rigid overlap	non rigid overlap
1	76.37	94.76
2	74.21	93.35
3	74.92	95.42
4	75.10	94.37
5	73.81	82.21
6	77.49	94.59
7	72.97	91.27
8	75.33	80.52
9	75.42	92.57
10	75.96	92.18
mean	**75.16**	**91.12**

Table 7.3: Overlap measure for 2D/3D registration (in percent)

orthogonal turntable-angles using a prototype of an interventional X-ray imaging system equipped with the digital X-ray detector described in section 3.3.2 (Figure 3.12). Preprocessing of the images was performed following section 3.3.2. The source position was calculated as explained in section 3.4.1. The turntable axis was was determined using a dedicated calibration algorithm. Prior to the experiment a CT-scan of the cadaver was acquired (voxel size: $0.273438 \times 0.273438 \times 1.0mm^3$). The vertebra used for the experiment (L2) was segmented manually for later evaluation (see below). L2 was segmented likewise manually in the two orthogonal X-ray images, see Fig. 7.12 for the lateral view. Subsequently the extracted contours were registered with the vertebra shape model using all deformation modes. Table 7.4 shows the results for this experiment. The experiments with simulated data showed that the rigid parameters can always be recovered reliably using one lateral and one AP view. Taking this into account, allows to evaluate the result of the non rigid registration by rigidly registering the deformed model with the segmented CT-model (reference) of the vertebra. This procedure supersedes otherwise necessary fiducial based registration. There are at least several possible reasons for the poorer result with respect to the experiment carried out with simulated projective data. The outliers in the experiments using simulated data (cases 5 and 8) are in the same range as the result for real data. This indicates that the statistical model does not contain sufficient representative information about the encountered shape variations. Further the calibration procedure might have introduced geometrical errors and it must certainly be taken into account that the cadaver spine belongs to a 80 year old specimen; the vertebrae show heavy degenerative changes. The computed shape model was based on a younger population, thus such shape variations occurring in older specimen are not gathered by the model.

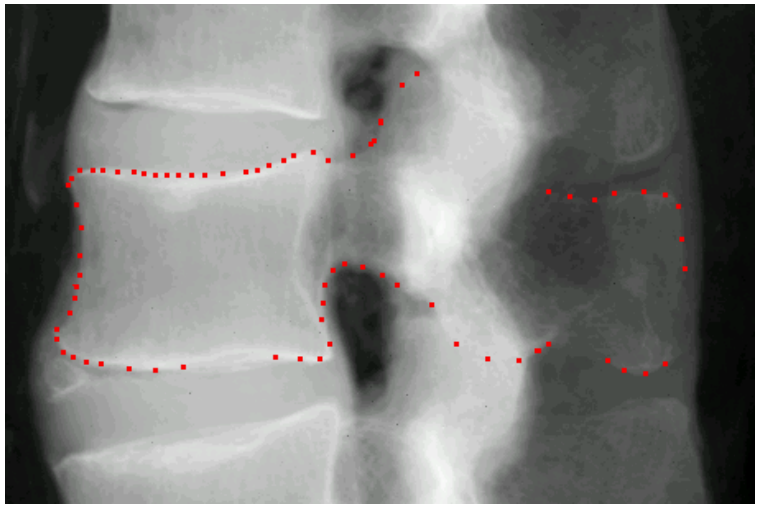


Figure 7.12: Manually segmented contour points on a lateral X-ray image of a cadaver spine (segmented points are enlarged for visualization).

rigid RMS model-projections	non rigid RMS model-projections	rigid RMS model-reference	non rigid RMS model-reference
1.43	0.83	1.52	1.27

Table 7.4: Matching errors for real data (in mm)

7.5 Model based Segmentation of X-ray images

The above experiments were carried out based on manually segmented X-ray images. As said in the introduction, this issue (though very important for the practical relevance of the method) is not of primary interest in this work. However, this section presents some preliminary qualitative results obtained with a model based approach combining low level pre-segmentation with knowledge based subsequent high-level segmentation, taking into account the segmentation-registration duality ([Ion98]).

In a first step low-level pre-segmentation is performed using a Canny-Deriche filter with subsequent thresholding, resulting in an oversegmented image. False positives are then detected by establishing correspondence between all detected contour pixels and the actual contour generators of the shape model. After projection of the contour generators into the image plane using current projection parameters, a deformation grid is associated to the contours (Figure 7.13). Each node of the grid is associated to displacement vector $V_{i,j}; (i,j) \in [0, I_x] \times [0, I_y]$, where $I_x(I_y)$, is the number of horizontal (vertical) nodes. Thus each point P of all contour points can be locally deformed by bi-linear interpolation of the positions of the four nodes the point P is associated to, $(V_{i,j}, V_{i+1,j}, V_{i,j+1}, V_{i+1,j+1})$, followed by a rigid transformation $T(\alpha; t_x, t_y)$ where α is the rotation angle and (t_x, t_y) the translation.

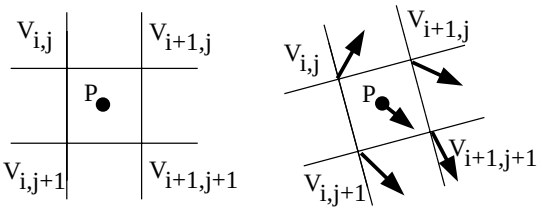


Figure 7.13: Grid deformation: The displacements are represented by the vectors associated to the grid nodes.

The objective function used to determine the optimal parameters for the grid displacement comprises three different components. The first term represents the distance from the transformed point P_i to the nearest point Q_j in the pre-segmented image, the second term regularizes the dimension of the vectors $V_{i,j}$ associated with the (i,j) grid node and the third term regularizes the similarity of vectors associated with neighboring nodes. To speed up the distance computations, a precomputed distance map is used.

Figure 7.14 shows different stages of this method using a multi-resolution approach. The projected contours of the model (yellow) get closer to the contour pixel obtained with the Canny filter (white) for higher grid levels. The different gray levels in the background represent the different distance values that are pre-computed and stored in a distance map to speed up the optimization. At the final resolution, only those pixels in the image that were matched with a contour generator of the model are retained for the 2D/3D registration. Assuming a sufficient accurate initialization of the rigid parameters for the model, the idea of the overall approach is to iterate this process until convergence.

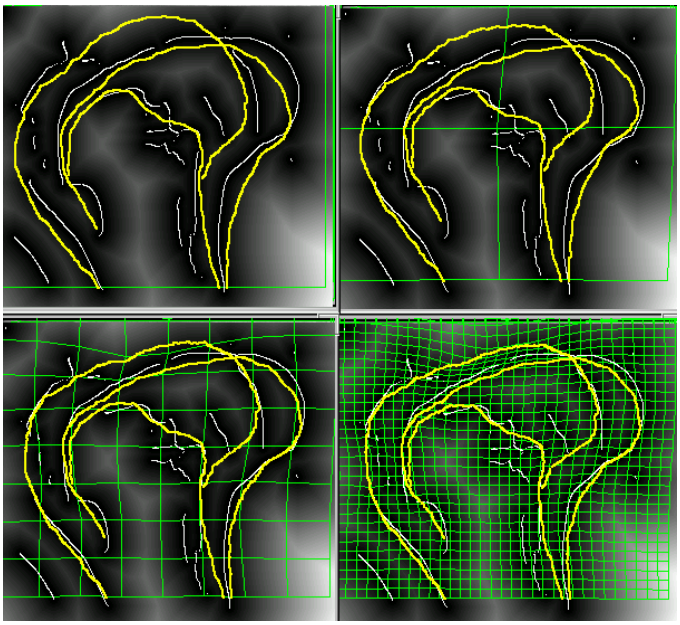


Figure 7.14: 2D image segmentation using model guided multi-grid deformation

7.6 Hybrid Matching

In chapter 6 it was shown that the shape of a femur can be reconstructed with sufficient accuracy for certain applications using relatively few scattered point data acquired in a highly restricted surface area. However, due to insufficient geometrical constraints (depending on the shape and on the spatial distribution of the acquired points), it might be mandatory to provide the correct pose parameters prior to the non rigid matching.

In [ea01b] authors report that the intra-operative rigid registration of a point cloud that has been acquired on a restricted dorsal surface area of a vertebra with a pre-operative acquired CT-model, can lead to unstable results (rotational uncertainty around the transversal axis) due to insufficient geometrical constraints in the accessible surface area of the vertebra. It is interesting to investigate if an additional X-ray image could provide sufficient extra information to sufficiently constrain the solution. This results in combining the methods proposed in this and the previous chapter, constituting a hybrid registration approach relying on both 2D projective and 3D point data. This is not only interesting for rigid registration but also for the more difficult problem of non rigid registration.

In this section it is investigated how far it is possible to reconstruct the shape of a vertebra using few 3D point data plus one lateral X-ray view. Using a single X-ray image alone does not provide sufficient information for the 2D/3D registration algorithm: Using a lateral view for instance (see Figure 7.15), doesn't provide a stable solution for the parameter determining the translation in the indicated (AP) direction. The instability increases the more parallel the beam becomes (i.e for greater source-detector distances).

An experiment was carried out using approximately 70 data points acquired on a CT

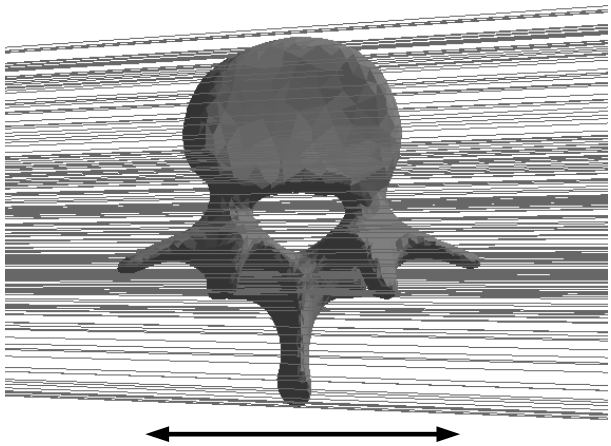


Figure 7.15: If only a single X-ray image is acquired there is no stable solution for the translation in the indicated direction.

based model in a small area accessible during spine surgery thus simulating the situation encountered in CAS systems where rigid registration with a pre-operative CT-scan is performed (see section 2.3.1). One lateral X-ray shot was then taken from the same model (see Fig. 7.16). Subsequently the shape model was fitted to both data - the 3D points and the segmented contours in the X-ray image - simultaneously. The results for this experiment are summarized in table 7.5. They compare well to the results for pure 2D/3D registration using two orthogonal X-ray views presented in section 7.4.1.

rigid RMS model-(projections,data points)	non rigid RMS model-projections	rigid RMS model-reference	non rigid RMS model-reference
1.18	0.40	1.88	0.68

Table 7.5: Matching errors for hybrid Matching using 70 data points and one lateral X-ray image (in mm)

7.7 Conclusion

Statistical shape models have been proven to be effective for different tasks in the field of computer vision such as segmentation of 2D and 3D images. While the previous chapter presented a new approach for nonrigid 3D/3D registration using such a model, this chapter investigated a new approach to perform nonrigid 3D/2D registration between such a model and very few calibrated X-ray images. Contour based registration algorithms suffer from the potential drawback that their accuracy directly depends on the correct segmentation of the objects contour in the image. The presented approach is relatively robust with respect to this problem in such a way that good matching results are obtained even if the object can be segmented only partially in the X-ray images. Computation time

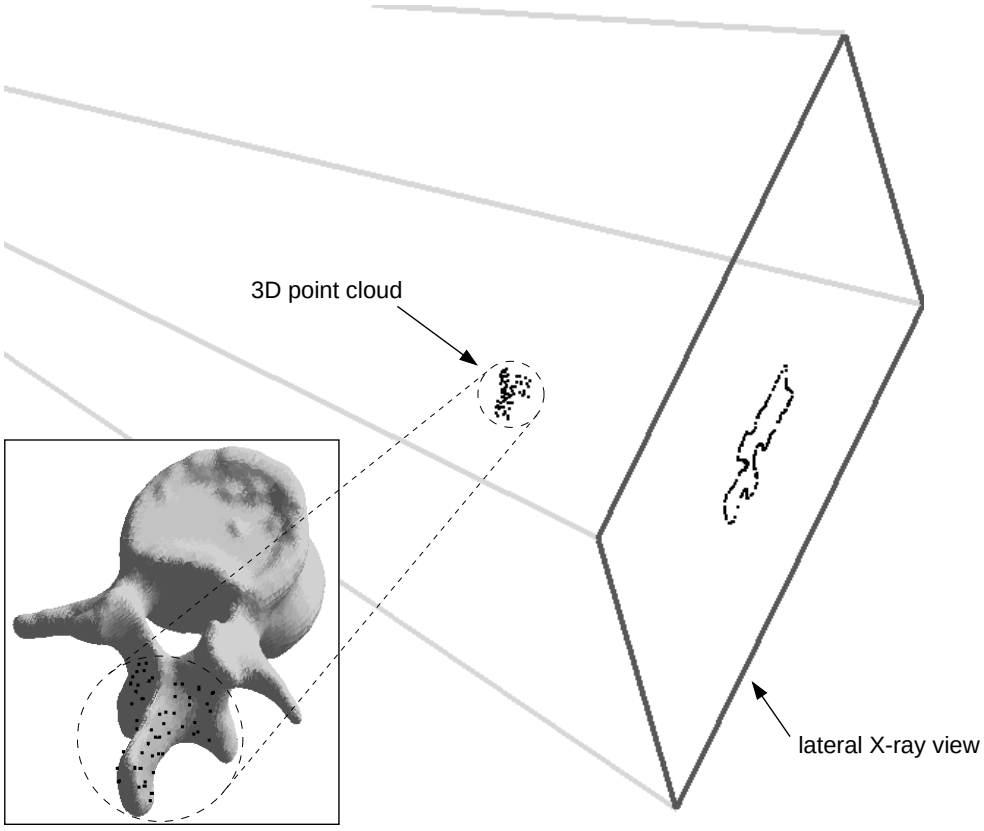


Figure 7.16: Hybrid matching using one lateral X-ray projection plus 70 points acquired in a small area accessible during spine surgery.

of the current implementation directly depends on the number of used projection rays and is less than one minute on a standard workstation when using two images with a total number of 400 projection rays and 30 deformation modes. A first experiment based on real data acquired with a digital X-ray detector has been presented as well and showed encouraging first results. However further work has specially to be addressed to the problem of reliable automatic segmentation of a certain portion of contour points. First qualitative results using a model based 2D/2D segmentation algorithm taking into account the duality between segmentation and registration are encouraging but further investigation must show, whether accuracy and especially robustness are sufficient for the given task. Finally, combining 3D and 2D data in a hybrid approach to recover shape seems very promising, but again further investigation and experiments are necessary to study how many data points at what locations and how many X-ray projections are necessary to allow reliable reconstruction.

Chapter 8

Discussion and perspectives

This dissertation focused on the development of new methods for the reconstruction of 3D anatomical surface models for multiple purposes such as surgical planning and visualization relying on intra-operative data only, more precisely on incomplete point data and two or three calibrated X-ray projections. The proposed approach consists in matching the data with a statistical deformable shape model thus incorporating a priori knowledge into the reconstruction process. The computation of such a statistical model is a non-trivial task due to the necessary segmentation and registration of shapes in a potentially large database. A method based on a generic model of the object is proposed to segment training shapes and to establish point to point correspondence simultaneously in a set of CT images. A volumetric coarse to fine deformation method based on free form deformations is used to match the generic model to the image data. Two methods have been investigated to attract the model boundary to the object contours in the gray-level image. The first method is based on the bone threshold in the gray level image. Experiments performed on CT images of the vertebral column have shown that good segmentation results can be obtained except for regions where two neighboring vertebrae are articulating and thus the bone gaps are too narrow. A second approach based on a statistical gray-level model for each model vertex leads to similar results.

The computed shape model can be matched to scattered point data using a non rigid 3D/3D registration algorithm, based on a least square fitting. The application of this method for intra and extrapolation of sparse point data is demonstrated within a system for computer assisted reconstruction of the anterior cruciate ligament using simulated and real data.

To reconstruct a surface from few calibrated X-ray images the statistical shape model is matched to the object contours segmented on the calibrated images based on a new non rigid 3D/2D registration method. Experiments are performed on a statistical model of lumbar vertebrae based on a data base containing 30 specimens. Experimental results indicate that the obtainable resolution is sufficient for the given task. It is further shown that hybrid registration combining both 3D/3D and 3D/2D registration, might be a very interesting option for certain computer assisted surgery applications. Contour based registration algorithms suffer from the potential drawback that their accuracy directly depends on the correct segmentation of the object contours in the image. The presented approach

is relatively robust with respect to this problem in such a way that good matching results are obtained even if the object can be segmented only partially in the X-ray images. However further work has to be addressed specifically to the problem of reliable automatic segmentation of a sufficient portion of contour points. First qualitative results using a model based 2D/2D segmentation algorithm taking into account the duality between segmentation and registration are encouraging but further investigation must show, whether accuracy and especially robustness are sufficient for the given task. Finally, combining 3D and 2D data in a hybrid approach to recover shape seems very promising, but again further investigation and experiments are necessary to study how many data points at what locations and how many X-ray projections are necessary to allow reliable reconstruction.

One of the most obvious limitations of the approach is the fact that only shape whose expected variation can be captured by a statistical analysis of a population can be reconstructed; fractured organs can not be modeled. However, various interventions could benefit from such methods. In the case of reconstruction of a torn cruciate ligament for instance, there is usually no pathologic shape variation of the adjacent bones (tibia, femur). Considering pedicle screw placement for spine instrumentation in the case of a vertebra compression fracture for instance, the shape of the vertebrae the screws are attached to (which are adjacent to the fractured one), is not pathologic; at least there is no pathologic shape variation associated to the reason for the surgery - the fracture.

8.1 Future Work

8.1.1 Automatic Shape extraction and landmarking

The presented algorithm for automatic shape extraction from CT images shows promising results. As the primary goal was to reconstruct surfaces, it has not been investigated quantitatively how reliably landmarks or semi-landmarks respectively are generated.

Tasks going further than reconstruction of the surface only, may require very accurate landmark setting. Considering the clinical application of pedicle screw placement for instance it could be interesting to propose an ideal screw trajectory and/or the diameter of the screw to be inserted together with the surface reconstruction. These tasks rely on precise landmarking in contrast to the pure reconstruction task where gliding of landmarks tangential to the surface do not alter the reconstruction result, though would influence the computation of the ideal screw trajectory.

It would be interesting to perform a comparison between the models built with different methods. In order to carry out a quantitative comparison it is necessary to define a measure of model quality. The definition of such a measure is in itself an interesting issue. Obviously, different methods will yield different sets of landmarks which precludes a landmark based comparison. If one defines a given segmentation task, a comparison could be established on the basis of the segmentation accuracy. Although these measures can have a prominent practical value to determine the best model-building technique for a given problem, the conclusion will remain task-dependent. Possibly, other more task-independent criteria related to the compactness and generalizability of the built models could be within the interesting candidate measures to explore.

The type of model used throughout this thesis is a boundary model. An interesting alternative is to use volumetric models taking into account not only boundary information but also the gray level distribution (texture) inside the model. One example of such a model would be the 3D version of the Active Appearance Model [TT00]. However even with modern computers especially extreme memory requirements prohibit straightforward implementation. Initial work on extending AAMs to 3D has been done by Wolstenholme et al. [WT99], where it was shown that wavelet compression successfully could be used to reduce the memory requirements without noteworthy loss of quality. The texture PCA was simply performed on wavelet coefficients instead of the raw pixels. Reliable image interpretation was obtained at a compression ratio of 20:1.

To improve the robustness of the shape extraction algorithm one could use a bootstrap method and compute a preliminary statistical model of all already segmented objects of a given class and then use this temporary model to approximately segment a new image. Final finer shape adjustments could then be established using the FFDs.

8.1.2 Statistical Model

The set of shapes used to build the statistical model forms a cloud of points in the $3M$ dimensional space, which is thought of to be drawn from a *single* Gaussian probability distribution. This might be incorrect in the general case. For example one can imagine that the female femur shapes and the male femur shapes are each drawn from a single Gaussian distribution with different parameters. Clustering techniques have to be applied to study this problem. A general approach to overcome the problem is to use a density estimation technique, such as the kernel method [Sil86] representing the distribution as a sum of Gaussians, one placed at every original data point. To approximate the distribution further, a mixture of a small number of Gaussians can be used instead [MB88]. See [CT97] for further details of this method.

Instead of capturing the statistics of the vertices of the triangle meshes representing the shapes, one could apply the PCA directly to the vertices of the deformation octree spline, which is used to find corresponding points within the population. This would result in a volume-based model rather than a surface-based model.

8.1.3 Matching of the statistical model with intra-operative data

For the 3D/3D registration algorithm it is interesting to investigate the minimal number and the optimal spatial distribution of points necessary to reconstruct a surface with a given accuracy. Furthermore more work has to be done to evaluate the extrapolation power of the proposed method. In the presented clinical application the reconstructed shape has been used for visualization purposes. Thus millimetric or even submillimetric precision is not mandatory. However if the model shall be used to incorporate other criteria for optimal graft placement in the case of ACL surgery, a certain guaranteed precision will

become an issue.

The scattered point data are usually acquired with an optical localizer system but also could originate from other intra-operatively available sensors like laser scanners, or 2.5D ultrasound sensors. In this case the least square criterion must be replaced by a more robust criterion, since outliers are inevitable. Further investigations have to show whether the accuracy of the model is sufficient to outperform the bicubic spline in interpolating the data points in terms of a more accurately computed isometry map. The first experiments show that reconstruction from very limited data is possible using only a few modes of the model. This validates the statistical approach in the particular instance of the femur although it needs to be confirmed on a large population. Using this method, it needs to be investigated how many points, at which location are necessary for correct reconstruction.

Appendix A

Glossary of Terms

Computer Assisted Surgery Systems use methods and systems to help the surgeon or the physician use multimodality data (mainly medical images) in a rational and quantitative way, in order to plan but also to perform medical interventions through the use of passive, semi-active or active guiding systems.

Segmentation is the process of delineating and labeling image regions (of any dimensionality) as distinct structures.

Registration is the determination of a geometrical transformation or mapping between any two spaces in such a way that points in one space are aligned with corresponding (homologous) points in the other space.

Deformable Shape Models can be characterized as models, which under an implicit or explicit optimization criterion, deform a shape to match a known object in a given image.

Shape is the characteristic surface configuration of an object (outline, contour) and invariant to translation, rotation and scaling.

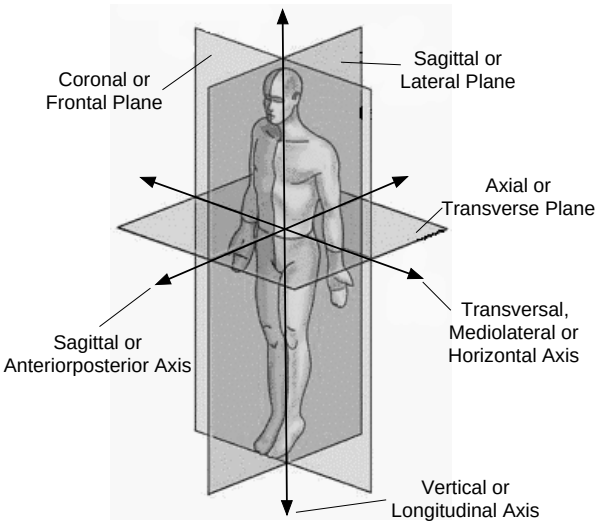
Anatomical Landmarks are salient and homologous points of the morphology of the visible anatomy and are accurately locatable for every example of a given object class and correspond between different specimens.

Semi-landmarks or pseudo-landmarks are points that are not necessarily determinable on a certain specimen alone but that correspond across all the specimens of a data set under a reasonable model of deformation from their common mean [Boo96]. Thus they can be used as if they were anatomical landmarks.

Appendix B

Anatomy

B.1 Anatomical terms for directions



Anatomic Terms	Direction
Medial	Toward the midline of the body
Lateral	Away from the midline of the body
Proximal	Toward a reference point (extremity)
Distal	Away from a reference point (extremity)
Inferior	Lower or below
Superior	Upper or above
Cranial	Head
Caudal	Tail, tail end
Anterior	Toward the front
Posterior	Toward the back
Dorsal	Posterior
Ventral	Anterior

Figure B.1: Anatomical reference system and anatomical terms for directions

B.2 Basic Bone Structure

Bones are organs composed of hard living tissue providing structural support to the body, its scaffolding. It is a hard matrix of calcium salts deposited around protein fibers. Minerals make bone rigid and proteins (collagen) provide strength and elasticity.

The outer layer of bone is called cortical bone. 80% percent of skeletal bone mass is cortical bone. Cancellous bone (also called trabecular bone) is an inner spongy structure that resembles honeycomb, which accounts for 20% of bone mass. The inner bone cavities contain bone marrow where red blood cells are produced.

B.3 The Vertebral Column

The vertebral column (or spinal column) (Figure B.2 c) extends from the skull to the pelvis and is made up of 33 individual bones termed vertebrae. The vertebrae are stacked on top of each other and are grouped under the names cervical, thoracic, lumbar, sacral, and coccygeal, according to the regions they occupy. Starting from the top there are seven in the cervical region, twelve in the thoracic, five in the lumbar, five in the sacral, and four in the coccygeal. The purpose of the vertebral column is to give the body its stability and structure. The spine supports the muscles, discs and the nerves of the back. In between the vertebrae of the column, there are the inter vertebral discs which act as a cushion for absorbing shocks and giving the spine its flexibility. Besides that they give room for the nerves to exit the spinal canal. The natural curves in the spine, kyphotic and lordotic, provide resistance and elasticity in distributing body weight and axial loads the spine is exposed to during movement.

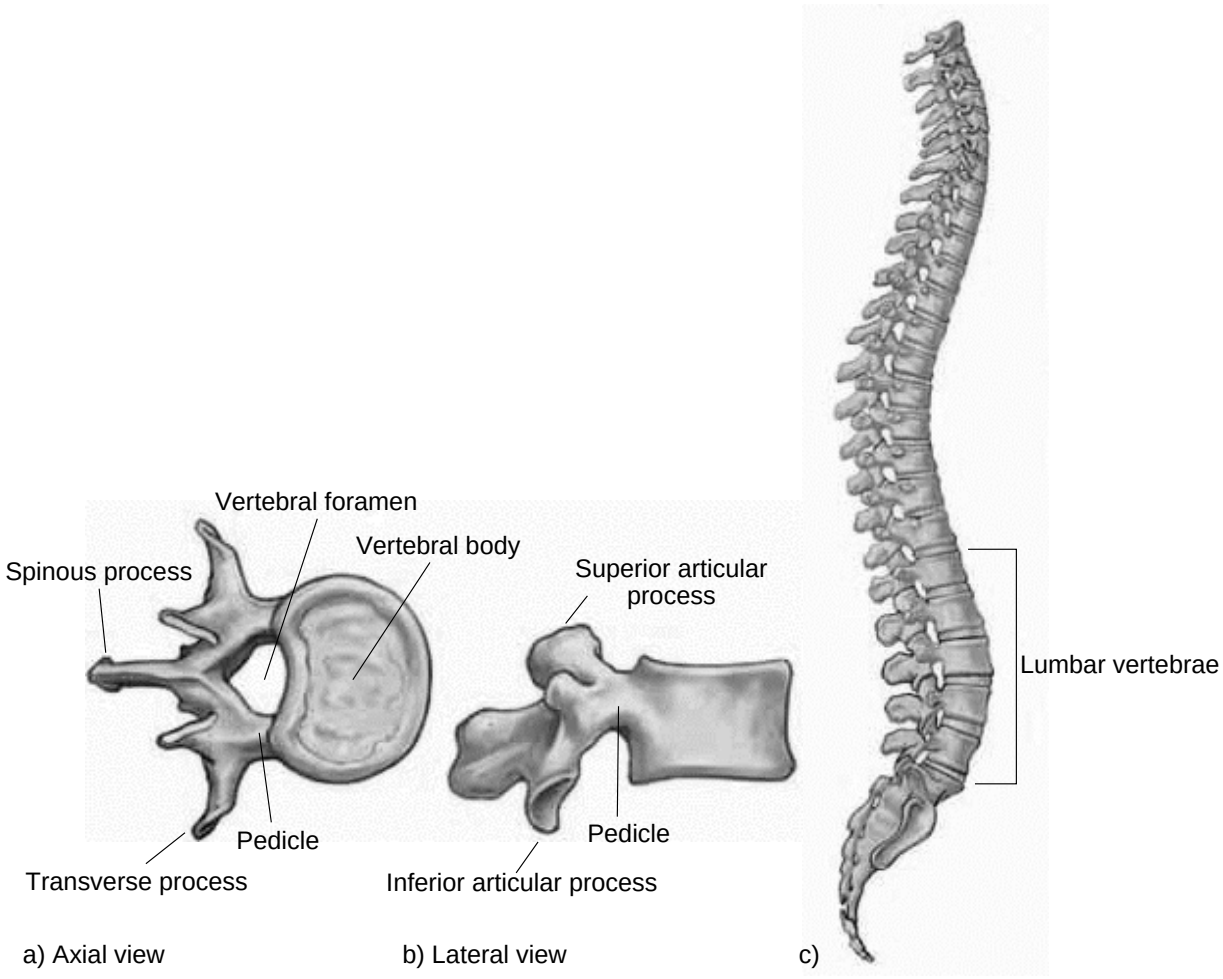


Figure B.2: Anatomy of the spine

A typical vertebra (see Figure B.2 a, b) consists of two essential parts: an anterior

segment, the body, and a posterior part, the vertebral or neural arch; these enclose a foramen, the vertebral foramen. The vertebral arch consists of a pair of pedicles and a pair of laminae, and supports seven processes: four articular, two transverse, and one spinous.

The pedicles are strong, cylindrical, anatomic bridges between the dorsal spinal elements and the vertebral body consisting of a strong shell of cortical bone and a core of cancellous bone.

B.4 The Knee Joint

The knee joint (Figure B.3) consists of three bones: the femur, the tibia and the patella. The two major articulations within the knee are the tibiofemoral and the patellofemoral joints. Motion of the tibiofemoral joint is complex. The knee does not move as a simple hinge but there is a complex combination of gliding and rolling of the femur on the tibia. The motion of the patellofemoral articulation is an up and down gliding on the front surface of the femur as the knee flexes and extends. The joint is cushioned by articular cartilage which covers the ends of the tibia and femur, as well as the underside of the patella. The lateral meniscus and the medial meniscus are cartilage pads which further cushion the joint, acting as shock absorbers between the bones. Ligaments help to stabilize the knee. The collateral ligaments run along the sides of the knee and limit sideways motion. The anterior cruciate ligament, or ACL, connects the tibia to the femur at the center of the knee. Its function is to limit rotation and forward motion of the tibia. The posterior cruciate ligament, or PCL limits backward motion of the tibia. For details about knee anatomy see for instance [JS92].

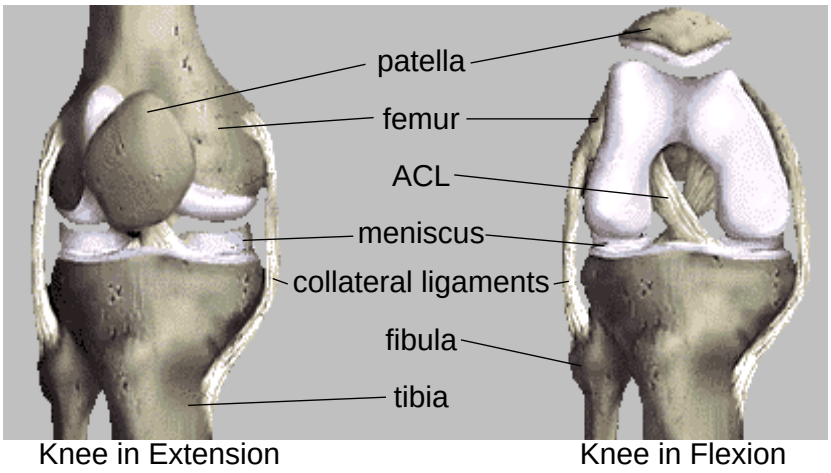


Figure B.3: Anatomy of the knee joint [sco]

Appendix C

Optimal rigid registration of two corresponding 3D point sets

In the following it is explained in detail, how to directly compute the optimal rotation $\mathbf{R}$ and translation $\mathbf{T}$ in the least square sense between a data point set and a model point set (taken from [BM92]). It is assumed that point correspondence between the data points and the model points has been established yet.

Assuming $\mathbf{q}$ is the unit quaternion $\mathbf{q}_R = [w \ x \ y \ z]^t$ describing the rotation and $\mathbf{q}_T = [q_4 q_5 q_6]^t$ a translation vector. The complete registration state vector can then be written as $\mathbf{q} = [\mathbf{q}_R|\mathbf{q}_T]^t$. $P = \{\mathbf{p}_i\}$ is the data point set to be aligned with the model point set $X = \{\mathbf{x}_i\}$, where the number of model points $N_x = N_p$ and where each point $\mathbf{p}_i$ corresponds to the point $\mathbf{x}_i$ with the same index. The mean square objective function to be minimized is

$$f(\mathbf{q}) = \frac{1}{N_p} \sum_{i=1}^{N_p} ||\mathbf{x}_i - \mathbf{R}(\mathbf{q}_R)\mathbf{p}_i - \mathbf{q}_T||^2.$$

The "center of mass" μ_p of the measured point set P and the center of mass μ_x for the X point set are given by

$$\mu_p = \frac{1}{N_p} \sum_{i=1}^{N_p} \mathbf{p}_i \; and \; \mu_x = \frac{1}{N_x} \sum_{i=1}^{N_x} \mathbf{x}_i.$$

The cross-covariance matrix $\sum_{px}$ of the sets P and X is given by

$$\sum_{px} = \frac{1}{N_p} \sum_{i=1}^{N_p} [(\mathbf{p}_i - \mu_p)(\mathbf{x}_i - \mu_x)^t] = \frac{1}{N_p} \sum_{i=1}^{N_p} [\mathbf{p}_i \mathbf{x}_i^t] - \mu_p \mu_x^t.$$

The cyclic components of the anti-symmetric matrix $A_{ij} = (\sum_{px} - \sum_{px}^T)_{ij}$ are used to form the column vector $\Delta = [A_{23} \ A_{31} \ A_{12}]^T$. This vector is then used to form the symmetric 4×4 matrix $Q(\sum px)$

$$Q(\sum px) = \begin{bmatrix} tr(\sum_{px}) & \Delta^T \\ \Delta & \sum_{px} + \sum_{px}^T - tr(\sum_{px})\mathbf{I}_3 \end{bmatrix}$$

where $\mathbf{I}_3$ is the 3×3 identity matrix. The unit eigenvector $\mathbf{q}_R = [w\ x\ y\ z]^t$ corresponding to the maximum eigenvalue of the matrix $Q(\sum px)$ is selected as the optimal rotation. The optimal translation vector is then given by

$$\mathbf{q}_T = \mu_x - \mathbf{R}(\mathbf{q}_R)\mu_p.$$

References

[ABG+97] P. Aglietti, R. Buzzi, F. Giron, A. Simeone, and G. Zaccherotti. Arthroscopic-assisted anterior cruciate ligament reconstruction with the central third patellar tendon. A 5-8 years follow-up. *Knee Surg, Sports Traumatol, Arthroscopy*, (5):138–144, 1997.

[ACC+96] N. Ayache, P. Cinquin, I. Cohen, L. Cohen, F. Leitner, and O. Monga. Segmentation of complex medical objects: a challenge and a requirement for computer assisted surgery planning and performing. In R. Taylor, S. Lavallee, G. Burdea, and R. Mosges, editors, *Computer Integrated Surgery*. MIT Press, Cambridge, MA, 1996.

[AS96] M. E. Algorri and F. Schmitt. Surface reconstruction from unstructured 3d data. *Comput. Graph. Forum*, 15(1):47–60, mar 1996.

[Aya91] N. Ayache. *Artificial Vision for Mobile Robots: Stereo Vision and Multisensory Perception*. MIT Press, Cambridge, Massachusetts, 1991.

[Bar81] A. H. Barr. Superquadrics and angle-preserving deformations. *IEEE Computer Graphics Applications*, 1:11–23, 1981.

[BBLRT99] X. L. Battel, Y. J. Bizais, C. Le Rest, and A. Turzo. Tomographic reconstruction using free-form deformation models. In *SPIE Medical Imaging Conference*, San Diego, California, USA, 1999.

[BCA94] E. Bardinet, L. Cohen, and N. Ayache. Fitting of iso-surfaces using superquadrics and free-form deformations. In *IEEE WEorkshop on Biomedical Image Analysis*, pages 184–193, Seattle, June 1994. IEEE Computer Society.

[Ben95] R. Benjamin. Object-based 3d x-ray imaging. In *CVRMed95*, pages 445–448, 1995.

[BK89a] R. Bajcsy and S. Kovacic. Multiresolution elastic matching. *Computer Vision, Graphics, and Image Processing*, 46:1–21, 1989.

[BK89b] R. Bajcsy and S. Kovacic. Multiresolution elastic matching. *Computer Vision, Graphics, and Image Processing*, 46(1):1–21, 1989.

[BLS93] E. Bittar, S. Lavallee, and R. Szeliski. A method for registering overlapping range images of arbitrarily shaped surfaces for 3-D object reconstruction. In *SPIE Vol. 2059, Sensor Fusion VI*, 1993.

[BM92] P.J. Besl and N.D. McKay. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.

[BM93] S. Beucher and F. Meyer. The morphological approach to segmentation: the watershed transformation. In E. R. Dougherty, editor, *Mathematical Morphology in Image Processing*, pages 433–481. Marcel Dekker, New York, 1993.

[Boi] J. Boissonat. Geometric structures for three-dimensional shape representation.

[Boo89] F. L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, June 1989.

[Boo96] F.L. Bookstein. Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. *Mecial Image Analysis*, 1(3):225–243, 1996.

[BT99] D. Brett and C.J Taylor. A framework for automated landmark generation for automated 3d statistical model construction. In *16th Conference on Information Processing in Medical Imaging, Visegrad, Hungary*, June 1999.

[BT00] D. Brett and C.J. Taylor. Construction of 3d shape models of femoral articular cartilage using harmonic maps. In *MICCAI'00*, 2000.

[BV91] R.M. Bolle and B.C. Vemuri. On three-dimensional surface reconstruction methods. *IEEE Trans PAMI*, 13(1):1–13, 1991.

[Can86] John Canny. A computational approach for edge detection. *IEEE/PAMI*, 8(6):679–697, november 1986.

[CCD+98] C. Chaussat, J. Chabbal, T. Ducourant, V. Spinnler, and G. Vieux. New superior detectivity csi/a-si 43cm x 43cm x-ray flat panel detector for general radiography provides immediate direct digital output and easy interfacing to digital radiographic systems. In H.U. Lemke, editor, *CAR*, 1998.

[CF93] L. Caponetti and A. Fanelli. Computer-aided simulation for bone surgery. *IEEE Computer Graphics & Applications*, 13(6):86–92, 1993.

[CL98] F. Chassat and S. Lavallee. Experimental protocol for accuracy evaluation of 6-D localizers for Computer Integrated Surgery : Application to four optical localizers. In *MICCAI'98*, pages 277–285, 1998.

[CLSC92] G. Champleboux, S. Lavallee, P. Sautot, and P. Cinquin. Accurate calibration of cameras and range imaging sensors, the NPBS method. In *IEEE Int. Conf. on Robotics and Automation*, pages 1552–1558, Nice France, May 1992.

[CMT98] F. Cordier and N. Magnenat-Thalman. Comparison of two techniques for organ reconstruction using visible human dataset. In *Visible Human Project Conference*, Bethesda, Maryland,USA, Oct.7-8 1998,.

[CRM94] G. Christensen, R. Rabbitt, and M. Miller. 3d brain mapping using a deformable neuroanatomy, 1994.

[CT97] T.F. Cootes and C.J. Taylor. A mixture model for representing shape variation. In *British Machine Vision Conference*, volume 3, pages 209–226. BMVA Press, 1997.

[CT98] A. Caunce and C.J Taylor. 3d point distribution models of the cortical sulci. In *Eighth International Conference for Computer Vision*, 1998.

[CTCG95a] 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, 1995.

[CTCG95b] 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, 1995.

[Der87] Rachid Deriche. Using Canny's Criteria to Derive a Recursively Implemented Optimal Edge Detector. *International Journal of Computer Vision*, pages 167–187, 1987.

[Des96] V. Dessenne. *GMCAO : applications a la ligamentoplastie et a la chirurgie orthognatique*. PhD thesis, University Joseph Fourier (Grenoble, France), October 1996.

[DKC+98] Frank Dachille, Kevin Kreeger, Baoquan Chen, Ingmar Bitter, and Arie Kaufman. High-quality volume rendering using texture mapping hardware. *1998 SIGGRAPH / Eurographics Workshop on Graphics Hardware*, pages 69–76, August 1998. Held in Lisbon, Portugal.

[DLJ+95] V. Dessenne, S. Lavallee, R. Julliard, P. Cinquin, and R. Orti. Computer assisted knee anterior cruciate ligament reconstruction : first clinical tests. In *Conference of Computer Vision, Virtual Reality, Robotics in Medicine (CVRMed'95), LNCS Series 905*, pages 476–480. Springer, 1995.

[DLO+95] V. Dessenne, S. Lavallee, R. Orti, R. Julliard, S. Martelli, and P. Cinquin. Computer assisted knee anterior cruciate ligament reconstruction : first clinical tests. *J. of Image Guided Surgery*, 1(1):59–64, 1995.

[ea] C.T. Mehlman et al. Radiation exposure to the orthopaedic surgical team during fluoroscopy: "How far away is far enough?", journal =.

[ea90] A. L. Yuille et al. Deformable templates for feature extraction from medical imageses. In O. Faugeras, editor, *First European Conference on Computer Vision*, volume 3, pages 413–417. Springer-Verlag, Berlin/New York, 1990.

[ea93] R. Sanders et al. Exposure of the orthopaedic surgeon to radiation. *Bone Joint Surg*, (75):326–330, 1993.

[ea94] E.E. Khalfayan et al. Relationship of tunnel placement and functional outcome after anterior cruciate ligament reconstruction. In *Amer. Academy of Orthop. Surg. Annual meeting New Orleans*, february 1994.

[ea97] R. Fahrig et al. Three-dimensional computed tomographic reconstruction using a c-arm mounted xrii: Correction of image intensifier distortion. *Med. Phys.*, 7(24):1097–1106, 1997.

[ea98] P. Mozzo et al. A new volumetric ct machine for dental imaging based on the cone-beam technique: preliminary results. *Eur. Radiol.*, 9(8):1558–1564, 1998.

[ea99a] A. Kelemen et al. Elastic model-based segmentation of 3d neuroradiological data sets. *IEEE Transactions on Medical Imaging*, 18(2):829–839, 1999.

[ea99b] J. Lotjonen et al. Model extraction from magnetic resonance volume data using the deformable pyramid. *Mecial Image Analysis*, 3(4):387–406, 1999.

[ea99c] P. Hellier et al. Medical image registration with robust multigrid techniques. In *Miccai*, 1999.

[ea00a] J. Montagnat et al. Surface simplex meshes for 3d medical image segmentation. In *ICRA*, 2000.

[ea00b] P.R. Andresen et al. Surface-bounded growth modeling applied to human mandibles. *IEEE Transactions on Medical Imaging*, 19(11):1053–1063, 2000.

[ea00c] Y. Wang et al. Shape-based 3d surface correspondence using geodesics and local geometry. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 644–651, 2000.

[ea01a] A.F. Frangi et al. Automatic 3d asm construction via atlas-based landmarking and volumetric elastic registration. In *Proc. Information Processing in Medical imaging (IPMI'01)*, pages 78–91, 2001.

[ea01b] C. Huberson et al. Surgical navigation for spine: Ct virtual imagery versus virtual fluoroscopy about 223 pedicle screws, in 88 patients. In *CAOSUSA*, pages 203–205, 2001.

[ea01c] D. Rueckert et al. Automatic construction of 3d statistical deformation models using non-rigid registration. In *MICCAI'01*, 2001.

[ECNM91] A.C. Evans, D. Collins, L. Neelin, and P. Marett. Warping of a computerized 3d atlas to match brain images volumes for quantitative neuroanatomical and functional analysis. *Image Processing*, (1445):236–246, 1991.

[Eea00] E. Euler and Wirth S. et al. 3d-imaging with an isocentric mobile c-arm. *Electromedica*, 68(2):122–126, 2000.

[Ela01] V. Elangovan. A direct method for segmenting tomographic data. Master's thesis, University of Utah, School of Computer Science, 2001.

[EM94] Herbert Edelsbrunner and Ernst P. Mücke. Three-dimensional alpha shapes. *ACM Transactions on Graphics*, 13(1):43–72, 1994.

[ER01] V. Elangovan and Whitakaker R. From sinograms to surfaces: A direct approach to the segmentation of tomographic data. In *MICCAI*, 2001.

[FA96] J. Feldmar and N. Ayache. Rigid, affine and locally affine registration of free-form surfaces. *IJCV*, 18(2):99–119, May 1996.

[FAB94] J. Feldmar, N. Ayache, and F. Betting. 3D-2D projective registration of free-form curves and surfaces. Technical Report 2434, INRIA, France, 1994.

[FBF77] J. H. Friedman, J.L. Bentley, and R. A. Finkel. An algorithm for finding best matches in logarithmic expected time. *ACM Trans. Math. Software*, 3(3):209–226, Sept. 1977.

[FH86] O.D. Faugeras and M. Hebert. The representation, recognition and locating of 3D objects. *Int. J. Robotic Res.*, 5(3):27–52, June 1986.

[FJ97] C.B. Frank and D.W. Jackson. The science of reconstruction of the anterior cruciate ligament. *J Bone Joint Surg*, 79A(10):1556–1566, 1997.

[FKH+01] A. Faridani, Buglione K., P. Huabsomboon, O. D. Iancu, and J. McGrath. *Introduction to Local Tomography*. Prentice-Hall, Englewood Cliffs, NJ, 2001.

[FL98] M. Fleute and Stephane Lavallee. Building a Complete Surface Model from Sparse Data Using Statistical Shape Models: Application to Computer Assisted Knee Surgery. In W. M. Wells, A. Colchester, and S. Delp, editors, *Medical Image Computing and Computer-Assisted Intervention-MICCAI'98*, pages 880–887. Springer Verlag, October 1998.

[FLS+97] J. Furia, D. Lintner, P. Saiz, H. Kohl, and P. Noble. Isometry measurements in the knee with the anterior cruciate ligament intact, sectioned and reconstructed. *Am J Sports Med*, 25(3):346–352, 1997.

REFERENCES

[GBB96] L. Gottesfeld Brown and T. Boult. Registration of planar film radiographs with computed tomography. In *IEEE Proceedings of MMBIA*, 1996.

[GDT+96] K.J. Gilhuijs, K. Drukker, A. Touw, P.J. van de Ven, and M. Van Herk. Interactive three-dimensional patient set-up in radiation therapy using digital portal images and computed tomography data. *Int. J. Radiation Oncology, Biol. Phys.*, 34(4):873–885, 1996.

[GH97] M. Garland and P. Heckbert. Surface simplification using quadric error metrics. In *SIGGRAPH'97*, 1997.

[GM97] G. Guy and G. Medioni. Inference of surfaces, 3D curves, and junctions from sparse, noisy, 3D data. *IEEE Trans PAMI*, 19(11):1265–1277, 1997.

[GM98] U. Grenander and M.I. Miller. Computational anatomy: An emerging discipline. Technical report, Center for Image Science (CIS), Brown University, 1998.

[GU97] Grevera G. and J. Udupa. An objective comparison of interpolation methods, 1997.

[Gue98] A. Gueziec. Anatomy-based registration of ct-scan and intraoperative x-ray images for guiding a surgical robot. *IEEE Transactions on Medical Imaging*, 17(5):715–728, October 1998.

[GW92] R. Gonzalez and R. Woods. *Digital Image Processing.* Addison-Wesley, third edition, 1992.

[HB80] W. Huck and U. Baer. A new targeting for stereotaxis procedures within the ct scanner. *Neuroradiology*, 19:13–17, 1980.

[HB95] S. Howell and S. Barad. Knee extension and its relationship to the slope of the intercondylar roof. *Am J Sports Med*, 23(3):288–294, 1995.

[HDDW92] H. Hoppe, T. DeRose, T.and McDonald J. Duchamp, and Stuetzle W. Surface reconstruction from unorganized points. In Catmull E. E., editor, *Computer Graphics (SIGGRAPH '92 Proceedings)*, pages 71–78, July 1992.

[HG97] P. Heckbert and M. Garland. Survey of polygonal surface simplification algorithms. Technical report, Carnegie Mellon University, USA, 1997.

[HL96] H. Horace and Y. Lijun. Constructing a 3d individualized head model from two orthogonal views. *The Visual Computer*, 12:254–266, 1996.

[HLC98] A Hamadeh, S Lavalle, and Ph Cinquin. Automated 3D CT and fluoroscopic image registration. *Journal of Imag Guided Surgery*, 1998.

[HM93] R. Horaud and O. Monga. *Vision par Ordinateur.* Hermes, Paris, 1993.

[Hof97] R. Hofstetter. Fluoroscopy based surgical navigation-concept and clinical applications. In H.U. Lemke, editor, *CAR*, 1997.

[Hor87] B. K. P. Horn. Closed-form solution of absolute orientation using unit quaternions. *J. Opt. Soc. Amer.*, 4(4):629–642, 1987.

[HR93] D. Huttenlocher and W. Rucklidge. A multi-resolution technique for comparing images using the Hausdorff distance. In *Proc. of the IEEE Computer Vision and Pattern Recognition Conf.*, pages 705–706, 1993.

[HRRS86] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A Stahel. *Robust Statistics: The Approach Based on Influence Functions.* Wiley, New York, 1986.

[HSLC95] A. Hamadeh, P. Sautot, S. Lavallee, and P. Cinquin. Towards automatic registration between CT and X-ray images : cooperation between 3D/2D registration and 2D edge detection. In *Second Symposium on Medical Robotics and Computer Assisted Surgery Proc. (MRCAS'95)*, pages 39–46, Baltimore, MA, nov. 1995. Wiley.

[HT93] S. Howell and M. Taylor. Failure of reconstruction of the anterior cruciate ligament due to impigement by the intercondylar roof. *J Bone Joint Surg*, 75A(7):1044–1055, 1993.

[Hub81] P. J. Huber. *Robust Statistics.* John Wiley and Sons, New York, New York, 1981.

[Ion98] G. Ionescu. *Segmentation et recalage d'images echographiques par utilisation de connaissances physiologiques et morphologiques.* PhD thesis, TIMC laboratory, Joseph Fourier University, Genoble, december 1998.

[JB94] F.A. Jolesz and S.M. Blumenfeld. Interventional use of magnetic resonance imaging. *Magnetic Resonance Quarterly*, (10):85–96, 1994.

[JC92] Shen J. and S. Castan. An optimal linear operator for step edge detection. *Computer Vision, Graphics, and Image Processing. Graphical Models and Image Processing*, 54(2):112–133, 1992.

[JLD98] R. Julliard, S. Lavallee, and V. Dessenne. Computer Assisted Anterior Cruciate Ligament Reconstruction. *Clinical Orthopaedics and Related Research*, (354):57–64, September 1998.

[JMC+93] J. Jerosch, J. Malms, W.H. Castro, R. Wagner, and L. Wiesner. Control of pedicle screw placement following instrumented posterior lumbar fusion (in German). *Z Orthop*, (130):479–483, 1993.

[JR93] J.J. Jacq and C. Roux. Automatic registration of 3D images using a simple genetic algorithm with a stochastic performance function. In *IEEE Engineering in Medicine and Biology Society Proceedings*, pages 126–127, San Diego, CA, October 1993.

[JRH92] H. Jiang, R.A. Robb, and K.S. Holton. A new approach to 3-D registration of multimodality medical images by surface matching. In *Visualization in Biomedical Computing (VBC'92), SPIE vol. 1808*, pages 196–213, 1992.

[JS92] R.P. Jacob and H.U. Staubli. *The Knee and the Cruciate Ligaments.* Springer-Verlag, 1992.

[KCD98] M. Kamath, S. Chaudhuri, and U.B. Desai. Direct parametric object detection in tomographic-images. *IVC*, 16(9-10):669–676, July 1998.

[KHJH88] Y.S. Kwoh, J. Hou, E.A. Jonckheere, and S. Hayati. A robot with improved absolute positioning accuracy for CT guided stereotactic brain surgery. *IEEE Transactions on Biomedical Engineering*, 35(2):153–160, 1988.

[Kit96] Y. Kita. Elastic-model driven analysis of several views of a deformable cylindrical object. *PAMI*, 18(12):1150–1162, December 1996.

[Kli90] Catherine Klifa. 3d reconstructing using a limited number of projections. In *SPIE Visual Communications and Image Processing*, 1990.

[KS87] A. Kak and M. Slaney. *Principles of Computerized Tomographic Imaging.* IEEE Press, New York, 1987.

[KT98] A.C.W. Kotcheff and J.C. Taylor. Automatic construction of eigenshape models by direct optimization. *Medical Image Analysis*, 2(4):303–314, 1998.

[KWT88] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1:321–331, 1988.

[Lau95] A. Laurentini. How far 3d shapes can be understood from 2d silhouettes. *PAMI*, 17(2):188–195, February 1995.

[Lav96] S. Lavallee. Registration for Computer Integrated Surgery : methodology, state of the art. In R. Taylor, S. Lavallee, G. Burdea, and R. Mosges, editors, *Computer Integrated Surgery.* MIT Press, Cambridge, MA, 1996.

[LBC+99] S. Lavallee, E. Bittar, B. Couteau, Y. Payan, and M.C. Hobatho. Elastic registration of 3D images and inferrence using octree-splines and image radient features. In *Proceedings of the International Workshop on Biomedical Image Registration*, Bled, Slovenia, August 1999.

[LBM+96] L. Launay, P. Bouchet, E. Maurincomme, M. Berger, and J. Mallet. A flexible iterative method for 3d reconstruction from x-ray projections. In *Proceedings of ICPR*, 1996.

[LCT97] S. Lavallee, P. Cinquin, and J. Troccaz. Computer Integrated Surgery and Therapy: State of the Art. In C. Roux and J.L. Coatrieux, editors, *Contemporary Perspectives in Three-Dimensional Biomedical Imaging*, chapter 10, pages 239–310. IOS Press, Amsterdam, NL, 1997.

[Lel99] B. Lelieveldt. *Anatomical models in cardiovasculair image analysis.* PhD thesis, University of Leiden, october 1999.

[LFK94] L. Lemieux, D.R. Fish, and N.D. Kitchen. A patient-to-computed-tomography image registration method based on digitally reconstructed radiographs. *Medical Physics*, november 1994.

[LK00] C. Lorenz and N. Krahnstover. Generation of point-based 3d statistical shape models for anatomical objects. *Computer Vision and Image Understanding*, 77:175–191, 2000.

[LL94] P. Lacroute and M. Levoy. Fast volume rendering using a shear-warp factorization of the viewing transformation. *Proceedings of SIGGRAPH 94*, pages 451–458, July 1994. ISBN 0-89791-667-0. Held in Orlando, Florida.

[LMNK99] J. Lotjonen, E. Magnin, J. Nenonen, and T. Katila. Reconstruction of 3d geometry using 2d profiles and a geometric prior model. *IEEE Trans. Medical Imaging*, 18(10):992–1002, 1999.

[Lon98] S Loncaric. A survey of shape analysis techniques. *Pattern Recognition*, 31(8):983–1001, 1998.

[Lor87] Lorensen. Marching cube: a High Resolution 3D Surface Construction Algorithm. *Computer Graphics, SIGGRAPH'87*, 21(4):163–169, 1987.

[LS95] S. Lavallee and R. Szeliski. Recovering the position and orientation of free-form objects from image contours using 3-D distance maps. *IEEE PAMI (Pattern Analysis and Machine Intelligence)*, 17(4):378–390, 1995.

[LSB91] S. Lavallee, R. Szeliski, and L. Brunie. Matching 3-D smooth surfaces with their 2-D projections using 3-D distance maps. In *SPIE Vol. 1570 Geometric Methods in Computer Vision*, pages 322–336, San Diego, CA, July 1991.

[LWM+94] J.T. Lea, D. Watkins, A. Mills, M.A. Peshkin, T.C. Kienzle, and S.D. Stulberg. Registration and immobilization for robot-assisted orthopaedic surgery. In *First Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'94)*, pages 63–68, Pittsburgh, Shadyside Hospital, 1994.

[MA83] W.N. Martin and J.K. Aggarwal. Volumetric descriptions of objects from multiple views. *PAMI*, 5(2):150–158, March 1983.

[MAS+91] S. Mirkovic, J.J. Abitbol, J. Steinmann, C.C. Edwards, and S.R. Garfin. Anatomic considerations for sacral screw placement. *Orthop Trans*, (15):235–236, 1991.

[MB88] G. Mclachlan and K.E. Basford. *Mixture Models: Inference and Application to Clustering.* Dekker, New York, 1988.

[MDSA00] J. Montagnat, H. Delingette, N. Scapel, and N. Ayache. Representation, shape, topology and evolution of deformable surfaces. application to 3d medical image segmentation. Technical Report RR-3954, INRIA, 2000.

[MEF99] B. Ma, R. E. Ellis, and D. J. Fleet. Spotlights: A robust method for surface-based registration in orthopedic surgery. In *MICCAI*, pages 936–944, 1999.

[mi3] mi3. http://mi3.vitamib.com/.

[MKG95] C. Morgan, V. Kalmam, and D. Grawl. isometry testing for anterior cruciate ligament reconstruction revisited. *Arthroscopy*, 11(6):647–659, 1995.

[MM97] R. Mencl and H. Muller. Interpolation and approximation of surfaces from threedimensional scattered data points. Technical Report 662, Universitaet Dortmund, December 1997.

[Mor95] C.D. Morgan. The all-inside ACL reconstruction. In: Operative Technique manual, Arthrex Inc., 1995.

[MT93] D. Metaxas and D. Terzopoulos. Shape and nonrigid motion estimation through physics-based synthesis. *IEEE Trans PAMI*, 15(6):580–591, 1993.

[MT96] T. McInerney and D. Terzopoulos. Deformable models in medical image analysis: a survey. *Medical Image Analysis*, 1(2):91–108, 1996.

[MV98] A. Maintz and M. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, 1998.

[NA92] R. Nastar and N. Ayache. Fast segmentation tracking and analysis of deformable objects. Technical report, INRIA, 1992.

[Nat86] F. Natterer. *The Mathematics of Computerized Tomography.* Wiley, 1986.

[Nik96] B. Nikkhahe. 3d reconstruction of the femoral bone using two x-ray images from orthogonal views. In H.U. Lemke, editor, *CAR*, 1996.

[Nik98] C. Nikou. Registration of mr/mr and mr/spect brain images by fast stochastic optimization of robust voxel similarity measures. *Neuroimage*, 8:30–43, 1998.

[OG98] C. Oblonsek and N. Guid. A fast surface-based procedure for object reconstruction from 3D scattered points. *Computer Vision and Image Understanding*, 69(2):185–195, 1998.

[PFTV92] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in C : The Art of Scientific Computing.* Cambridge University Press, Cambridge, England, second edition, 1992.

[Pic97] R. Pichumani. *Construction of a Three-dimensional Geometric Model for Segmentation and Visualization of Cervical Spine Images.* PhD thesis, Stanford University School of Medicine, 1997.

[PS91] A. Pentland and S. Sclaroff. Closed-Form Solutions for Physically Based Shape Modeling and Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(7):715–729, July 1991.

[PV97] R. Poli and G. Valli. Shape from radiological density. *Computer Vision and Image Understanding*, 65(3):361–381, March 1997.

[PWL+98] G. Penney, J. Weese, J. Little, P. Deesmedt, D. Hill, and D. Hawkes. A comparison of similarity measures for use in 2-d-3-d medical image reistration. *IEEE Trans. Medical Imaging*, 17:586–595, 1998.

[RBB+99] M. Roth, C. Brack, R. Burgkart, A. Czopf, H. Gotte, and A. Schweikar. Multi-view contourless registration of bone structures using a single calibrated x-ray fluoroscope. In *CARS*, 1999.

[RCSM86] R. Roy Camille, G. Saillant, and C. Mazel. Plating of thoracic, thoracolumbar and lumbar injuries with pedicle screw plates. *Orthop Clin North Am*, (17):147–159, 1986.

[Rot00] M. Roth. *Intraoperative fluoroskopiebasierte Patientenlageerkennung zur przisen Untersttzung chirurgischer Eingriffe.* PhD thesis, institute for computer science, Technical University, Munich, august 2000.

[Rou84] P. J. Rousseeuw. Least median of squares regression. *Journal of the American Statistical Association*, (79):871–880, 1984.

[RPSC99] H. Ray, H. Pfister, D. Silver, and T. Cook. Ray-casting architectures for volume visualization. *IEEE Transactions on Visualization and Computer Graphics*, 5(3):210–223, 1999.

[RTAR+95] S. Reboul, A. Taleb-Ahmed, M. Rousset, F. Wattrelot, and J. Dubus. Reconstruction d'objets binaires a partir de deux projections orthogonales par une technique inspiree de la theorie des graphes: la recherche du flot maximum a cout minimum. *Traitement du Signal*, (12)(2):326–341, 1995.

[Sai95] G. Saillant. *Complications de la vise pdiculaire, checs et complications de la chirurgie du rachis (in French).* Sauramps, Montpellier, 1995.

[sco] scoi. http://www.scoi.com.

[SD92] L.H. Staib and J.S. Duncan. Boundary finding with parametrically deformable models. *IEEE Trans. on PAMI*, 17(11):1061–1075, 1992.

[SdGD97] M. Sati, J.A. de Guise, and G. Drouin. Computer assisted knee surgery : diagnostics and planning of Knee Surgery. *Comp Aid Surg*, 2(2):108–123, 1997.

[SF93] T.O. Souryal and T.R. Freeman. Intercondylar notch size and anterior cruciate ligament injuries in athletes. *American Journal for Sports Medicine*, 21(4), 1993.

[SHK95] D.A. Simon, M. Hebert, and T. Kanade. Techniques for fast and accurate intrasurgical registration. *Journal of Image Guided Surgery*, (1):17–29, 1995.

[Sil86] B. Silvermann. *Density Estimation for Statistics and Data Analysis*. Chapmann and Hall, London, 1986.

[Sim93] E. Sim. Location of transpedicular screws fixation of the lower thoracic and lumbar spine. *Acta Othop Scand*, 64(1):28–32, 1993.

[Sim97] D. Simon. Intra-operative position sensing and tracking devices. In *Proceedings of the First Joint CVRMed / MRCAS Conference*, pages 62–64, June 1997.

[SKBG96] R. Szekely, A. Kelemen, C. Brechbuler, and G. Gerig. Segmentation of 2D and 3D objects from MRI volume data using constrained elastic deformations of flexible Fourier surface models. *Medical Image Analysis*, 1(1):19–34, 1996.

[SL96] R. Szeliski and S. Lavallee. Matching 3-D anatomical surfaces with non-rigid deformations using octree-splines. *Int. J. of Computer Vision (IJCV)*, (18)(2):171–186, 1996.

[SNP95] S. Sullivan, A. Noble, and J. Ponce. On reconstructing curved object boundaries from sparse sets of x-ray images. In *CVRMED*, 1995.

[SP86] T.W. Sederberg and S.R. Parry. Free-form deformations of solid geometric models. *Computer Graphics (SIGGRAPH'86)*, 20(4):151–160, 1986.

[STA96] G. Subsol, J.P. Thirion, and N. Ayache. Application of an automatically built 3D morphometric brain atlas: study of cerebral ventricle shape. In K.H. Hohne and R. Kikinis, editors, *Visualization in Biomedical Computing (VBC'96) Proc. LNCS 1131*, pages 373–382, Berlin, 1996. Springer-Verlag.

[TCCG92] Cootes T.F., Taylor C.J., D.H: Cooper, and J. Graham. Training models of shape from sets of examples. In *Proc. British Machine Vision Conference*, pages 266–275, Berlin, 1992. Springer.

[Ter88] D. Terzopoulos. The computation of visible-surface representations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-10(4):417–438, July 1988.

[TF88] D. Terzopoulos and K. Fleischer. Deformable models. *The Visual Computer*, 4(6):306–331, December 1988.

[Thi91] J.P. Thirion. A geometric alternative to computed tomography. Technical Report 1463, Inria, France, June 1991.

[TM91] D. Terzopoulos and D. Metaxas. Dynamic 3D models with local and global deformations: Deformable superquadrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(7):703–714, July 1991.

[TP91] M. Turk and A. P. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, 1991.

[Tsa86] R. Y. Tsai. An efficient and accurate camera calibration technic for 3D machine vision. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 68–75, Miami, 1986.

[TT00] Cootes T. and C.J. Taylor. Statistical models of appearance for computer vision, 2000.

[TWK88] D. Terzopoulos, A. Witkin, and M. Kass. Constraints on deformable models: Recovering 3D shape and nonrigid motion. *Artificial Intelligence*, 36:91–123, 1988.

[VD97] L. Vengrinovich Valery and Yuri B. Denkevich. Multi step 3d x-ray tomography for a limited number of projections and views. *Review of Progress in Quantitative Nondestructive Evaluation*, (16):317–323, 1997.

[VMH99] J. Vollmer, R. Mencl, and Mueller H. Improved laplacian smoothing of noisy surface meshes. In P. Brunet and R. Scopigno, editors, *Eurographics*, 1999.

[VRBa95] A. R. Vaccaro, S.J. Rozzolo, R.A. Balderston, and al. Placement of pedicle screws in the thoracic spine. *J Bone Joint Surg*, 77A(8):1200–1206, 1995.

[WE98] Rüdiger Westermann and Thomas Ertl. Efficiently using graphics hardware in volume rendering applications. *Proceedings of SIGGRAPH 98*, pages 169–178, July 1998. ISBN 0-89791-999-8. Held in Orlando, Florida.

[Wee99] J. Weese. Fast voxel-based 2d/3d registration algorthm using a volume rendering method based on the shear-warp factorization. In *SPIE*, 1999.

[WH96] P. Wunsch and G. Hirzinger. Registration of cad-models to images by iterative inverse perspective matching. In *International Conference on Pattern Recognition*, 1996.

[WSS+88] J.N. Weinstein, K.F. Spratt, D. Spengler, C. Brick, and S. Reid. Spinal pedicle fixation: reliability and validity of roentgenogram-based assessement and surgical factors on successful screw placement. *Spine*, 13:1012–1018, 1988.

[WT99] C. B. H. Wolstenholme and C. J. Talyor. Wavelet compression of active appearance models. In *Miccai*, 1999.

[YCH92] A. L. Yuille, D. S. Cohen, and P. Hallinan. Feature extraction from faces using deformable templates. *International Journal of Computer Vision*, (8):99–112, 1992.

[YDP92] N.C. Yaru, D.M. Daniel, and D. Penner. The effect of tibial attachment site on graft impingement in an anterior cruciate ligament reconstruction. *Am J Sports Med*, 20(2):217–223, 1992.

[YHC92] A. Yuille, P. Halinan, and D. Cohen. Feature extraction from faces using deformable templates. *Int. J. Comput. Vision*, 8:89–111, 1992.